

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
25 March 2004 (25.03.2004)

PCT

(10) International Publication Number  
**WO 2004/024161 A1**

(51) International Patent Classification<sup>7</sup>: **A61K 31/517**,  
C07D 403/04, A61P 3/06

(74) Agent: **SCHUCK, Alexander**; Isenbruck-Bösl-  
Hörschler-Wichmann-Huhn, Patentanwälte, Theodor-  
Heuss-Anlage 12, 68165 Mannheim (DE).

(21) International Application Number:  
PCT/EP2003/010036

(22) International Filing Date:  
10 September 2003 (10.09.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
02 020 255.2 10 September 2002 (10.09.2002) EP

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC,  
SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,  
SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM,  
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **PHENEX  
PHARMACEUTICALS AG** [DE/DE]; Im Neuenheimer  
Feld 515, 69120 Heidelberg (DE).

(72) Inventors; and

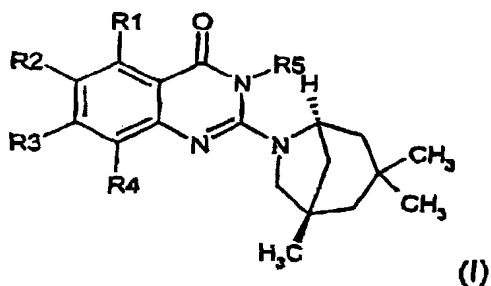
(75) Inventors/Applicants (*for US only*): **DEUSCHLE,  
Ulrich** [DE/DE]; Wundtstrasse 9/20, 69123 Heidelberg  
(DE). **KÖGL, Manfred** [AT/DE]; Hauptstrasse 131/4,  
69214 Eppelheim (DE). **BLUME, Beatrix** [DE/DE];  
Heinrich-Heine-Strasse 20, 69221 Dossenheim (DE). **AL-  
BERS, Michael** [DE/DE]; Obere Schulstrasse 3, 69221  
Dossenheim (DE). **KOBER, Ingo** [DE/DE]; Am Grossen  
Wald 30, 69251 Gaiberg (DE).

**Published:**

- with international search report
- before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: NOVEL 2-AMINO-4-OXOQUINAZOLONES AS LXR NUCLEAR RECEPTOR BINDING COMPOUNDS WITH  
PARTIAL AGONISTIC PROPERTIES



(57) Abstract: The present invention relates to compounds according to the general formula (I) which bind to the Liver X receptors (LXR receptors, LXRA $\alpha$ /NR1H3 and LXRB $\beta$ /NR1H2) and act as selective agonists of the LXR receptors. The invention further relates to the treatment of diseases and/or conditions through binding of said nuclear receptors and selective agonistic effects by said compounds and the production of medicaments using said compounds. In particular the compounds are useful in the treatment of hypercholesterolemia, obesity or other diseases associated with elevated lipoprotein (LDL) levels.

WO 2004/024161 A1

NOVEL 2-AMINO-4-OXOQUINAZOLONES AS LXR NUCLEAR RECEPTOR BINDING COMPOUNDS  
WITH PARTIAL AGONISTIC PROPERTIES

Liver X Receptor (LXR) is a prototypical type 2 nuclear receptor which activates  
10 genes upon binding to promoter region of target genes in a prototypical heterodimeric  
fashion with Retinoid X Receptor (hereinafter RXR, Forman et al., Cell, 81, 687-93,  
1995). The term LXR includes all subtypes of this receptor. Specifically LXR includes  
LXR $\alpha$  (also known as LXR $\alpha$ , RLD-1 and NR1H3) and LXR $\beta$  (also known as  
LXR $\beta$ , NER, NER1, UR, OR-1, R1P15 and NH1H2) and ligands of LXR should be  
15 understood to include ligands of LXR $\alpha$  or LXR $\beta$ . The relevant physiological ligands  
of LXR seem to be oxidized derivatives of cholesterol, including 22-  
hydroxycholesterol and 24,25(S)-epoxycholesterol (Lehmann, et al., Biol. Chem.  
272(6), 3137-40, 1997). The oxysterol ligands bound to LXR were found to regulate  
the expression of several genes that participate in cholesterol metabolism (Janowski,  
20 et al., Nature, 383, 728-31, 1996).

LXR is proposed to be a hepatic oxysterol sensor. Upon activation (e.g. binding of  
oxysterols) it influences the conversion of dietary cholesterol into bile acids by  
upregulating the transcription of key genes which are involved in bile acid synthesis  
25 such as CYP7A1. Hence, activation of LXR in the liver could result in an increased  
synthesis of bile acids from cholesterol which could lead to decreased levels of  
hepatic cholesterol. This proposed LXR function in hepatic cholesterol metabolism  
was experimentally confirmed using knockout mice. Mice lacking the receptor LXR $\alpha$   
lost their ability to respond normally to an increase in dietary cholesterol and did not  
30 induce transcription of the gene encoding CYP7A1. This resulted in accumulation of  
large quantities of cholesterol in the livers and impaired hepatic function (Peet, et al.,  
Cell, 93, 693-704, 1998).

Besides its important function in liver, LXR plays an important role in the regulation of  
35 cholesterol homeostasis in macrophages and intestinal mucosa cells where it  
upregulates cholesterol transporters from the ABC (=ATP binding cassette) family of  
membrane proteins (Repa, et al., J Biol Chem. 2002 May 24;277(21):18793-800).

5 These transporters are believed to be crucially involved in the uptake of cholesterol from the diet since mutations in their genes leads to diseases such as sitosterolemia (Berge, et al., Science (2000);290(5497):1771-5.).

10 Other members of the ABC transporter family seem to be responsible for the efflux of cholesterol from loaded macrophages, a process which is thought to prevent the generation of atherosclerotic lesions. Stimulation of LXR by synthetic ligands might result in an increased cholesterol efflux from macrophages and a decreased building up of cholesterol loaded atherosclerotic plaques (Venkateswaran, et al., PNAS (2000) 24;97(22):12097-102; Sparrow, et al., J Biol Chem (2002) 277(12):10021-7; 15 Joseph, et al., PNAS (2002);99(11):7604-9). Direct evidence that synthetic LXR ligands inhibit the development of atherosclerosis has been provided in two animal models of atherosclerosis: A significant reduction in the formation of atherosclerotic plaques were shown in two studies in animal models using full LXR agonists Joseph et al. PNAS (2002) 99:7604-9 and Terasaka et al. (2003) Terasaka et al. FEBS Lett. 20 (2003) 536:6-11. In addition, two recent reports have highlighted the potential use of LXR agonists in diabetes (Cao et al., (2003) J Biol Chem. 278:1131-6 and inflammatory disorders (Joseph et al., (2003) Nat Med. 9:213-9).

25 However, in animal studies it was observed that activation of LXR in the liver by full agonists like T0901317 does not only increase bile acid synthesis but also stimulates the de novo synthesis of fatty acids and triglycerids through the upregulation of key enzymes such as Fatty Acid Synthase (FAS) or Stearyl-CoA Desaturase (SCD-1) (Schultz, et al., Genes Dev (2000) 14(22):2831-8). Elevation of serum triglyceride levels is an indendent risk factor for atherosclerosis (for review see Miller (1999 ) 30 Hosp Pract (Off Ed) 34: 67-73.).

Thus, LXR activity needs to be selectively modulated for therapeutic benefit. In particular, compounds need to be found that stimulate reverse cholesterol transport, but do not significantly increase triglyceride levels. This might be particular relevant 35 for the usage of such compounds in diabetic patients since a even more severe lipogenic effect was reported for the full agonist T0901317 in db/db mice which serve as an animal model for diabetes (Chisholm et al. (2003) J.Lipid Res (epub August 16)).

5  
Therefore, an ideal synthetic LXR binding compound should have properties that retain the agonistic activity on hepatic bile acid formation and ABC-transporter - mediated decrease in cholesterol uptake from the diet and increased cholesterol efflux from macrophages. In parallel such a compound should lack the hyperlipidemic  
10 potential which is exerted through increased fatty acid and tricycleride synthesis.

To date only few compounds have been described which bind the LXR receptor and thus show utility for treating diseases or conditions which are due to or influenced by said nuclear receptor (Collins, et al., J Med Chem. (2002) 45(10):1963-6; Schultz, et  
15 al., Genes Dev (2000) 14(22):2831-8; Sparrow, et al., J Biol Chem (2002) 277(12):10021-7). No non-steroidal compounds have so far been described which show selectivity regarding the induction of ABC transporter genes without simultaneous induction of lipogenic genes like FAS and SREBP-1c (Kaneko et al. (2003) J Biol Chem (epub July 7).

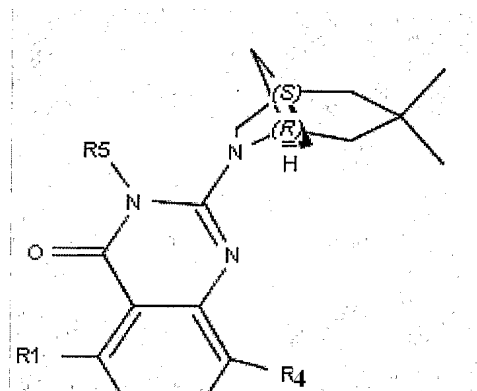
20 It is thus an object of the invention to provide for compounds which by means of binding the LXR receptor act as partial agonists of said receptor with a selective property regarding the upregulation of genes like the ABC transporters in macrophages and/or other cell types and a strongly reduced liability to increase the  
25 expression of genes involved in triglyceride synthetic pathways (like FAS and SREBP-1c). These compounds should show utility for treating diseases or conditions which are due to or influenced by said nuclear receptor.

It is further an object of the invention to provide for compounds that may be used for  
30 the manufacture of a medicament for the treatment of cholesterol associated conditions or diseases. It is still a further object of the invention to provide for compounds that lower serum cholesterol and/or increase High Density Lipoproteins (HDL) and/or decrease Low Density Lipoproteins (LDL). It is also an object of the invention to provide for compounds that may be used for the treatment of lipid  
35 disorders including hypercholesterolemia, atherosclerosis, Alzheimer's disease, skin disorders, inflammation, obesity and diabetes.

5 The present invention provides, *inter alia*, novel LXR nuclear receptor protein binding compounds according to the general formula (I) shown below. Said compounds are also binders of mammalian homologues of said receptor. Further the object of the invention was solved by providing for amongst the LXR nuclear receptor protein binding compounds according to the general formula (I) such compounds which act as partial agonists or mixed agonists / antagonists of the human LXR receptor or a mammalian homologue thereof. Further the object of the invention was solved by providing for amongst the LXR receptor protein binding compounds according to the general formula (I) such compounds which act as partial agonists of the human LXR receptor resulting therefore in the induction of ABC transporter proteins such as ABCA1 or ABCG1 in cell types such as macrophages but lacking a strong potential to induce genes involved in triglyceride synthetic pathways such as fatty acid synthase (FAS) or SREBP1c.

The invention provides for LXR agonists that may be used for the manufacture of a medicament for the treatment of cholesterol associated conditions or diseases. In a preferred embodiment compounds are provided that lower serum cholesterol and/or increase High Density lipoproteins (HDL) and/or decrease Low Density Lipoproteins (LDL). Also compounds are provided that may be used for the treatment of lipid disorders including hypercholesterolemia, atherosclerosis, Alzheimer's disease, skin disorders, inflammation, obesity and diabetes.

The invention provides for a compound of the formula (I), or pharmaceutical acceptable salts or solvates thereof, hereinafter also referred to as the "compounds according to the invention" including particular and preferred embodiments thereof.



5

(I)

The compounds of the invention can also exist as solvates and hydrates. Thus, these compounds may crystallize with, for example, waters of hydration, or one, a number  
 10 of, or any fraction thereof of molecules of the mother liquor solvent. The solvates and hydrates of such compounds are included within the scope of this invention.

In one embodiment of the invention in formula (I) above,  $R_1$ ,  $R_2$ ,  $R_3$ ,  $R_4$ , -  
 independent from each other - H, halogen, hydroxy, protected hydroxy, cyano, nitro,  
 15  $C_1$  to  $C_6$  alkyl,  $C_1$  to  $C_6$  substituted alkyl,  $C_1$  to  $C_7$  alkoxy,  $C_1$  to  $C_7$  substituted alkoxy,  
 $C_1$  to  $C_7$  acyl,  $C_1$  to  $C_7$  substituted acyl,  $C_1$  to  $C_7$  acyloxy, carboxy, protected carboxy,  
 carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl,  
 amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino,  
 (disubstituted)amino, carboxamide, protected carboxamide, N-( $C_1$  to  $C_6$   
 20 alkyl)carboxamide, protected N-( $C_1$  to  $C_6$  alkyl)carboxamide, N, N-di( $C_1$  to  $C_6$   
 alkyl)carboxamide, trifluoromethyl, N-(( $C_1$  to  $C_6$  alkyl)sulfonyl)amino, N-  
 (phenylsulfonyl)amino or phenyl,

and  $R_5$  is H,  $C_1$  to  $C_8$  alkyl,  $C_1$  to  $C_8$  substituted alkyl,  $C_7$  to  $C_{12}$  alkylphenyl or  $C_7$  to  
 25  $C_{12}$  substituted phenylalkyl.

In an other embodiment of the invention in formula (I) above  $R_1$ ,  $R_3$  and  $R_4$  are H,  $R_2$   
 is halogen, hydroxy, protected hydroxy, cyano, nitro,  $C_1$  to  $C_6$  alkyl,  $C_1$  to  $C_6$   
 substituted alkyl,  $C_1$  to  $C_7$  alkoxy,  $C_1$  to  $C_7$  substituted alkoxy,  $C_1$  to  $C_7$  acyl,  $C_1$  to  $C_7$   
 30 substituted acyl,  $C_1$  to  $C_7$  acyloxy, carboxy, protected carboxy, carboxymethyl,  
 protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected  
 amino, (monosubstituted)amino, protected (monosubstituted)amino,  
 (disubstituted)amino, carboxamide, protected carboxamide, N-( $C_1$  to  $C_6$   
 alkyl)carboxamide, protected N-( $C_1$  to  $C_6$  alkyl)carboxamide, N, N-di( $C_1$  to  $C_6$   
 35 alkyl)carboxamide, trifluoromethyl, N-(( $C_1$  to  $C_6$  alkyl)sulfonyl)amino, N-  
 (phenylsulfonyl)amino or phenyl,

and  $R_5$  is H,  $C_1$  to  $C_8$  alkyl,  $C_1$  to  $C_8$  substituted alkyl.

5

- The symbol "H" denotes a hydrogen atom.

The term "C<sub>1</sub> to C<sub>7</sub> acyl" encompasses groups such as formyl, acetyl, propionyl, butyryl, pentanoyl, pivaloyl, hexanoyl, heptanoyl, benzoyl and the like. Preferred acyl groups are acetyl and benzoyl.

The term "C<sub>1</sub> to C<sub>7</sub> substituted acyl" denotes the acyl group substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, cyclohexyl, naphthyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, guanidino, heterocyclic ring, substituted heterocyclic ring, imidazolyl, indolyl, pyrrolidinyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, nitro, C<sub>1</sub> to C<sub>6</sub> alkyl ester, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, N,N-di(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C<sub>1</sub> to C<sub>4</sub> alkylthio or C<sub>1</sub> to C<sub>4</sub> alkylsulfonyl groups. The substituted acyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

The term "substituted phenyl" specifies a phenyl group substituted with one or more, and preferably one or two, moieties chosen from the groups consisting of halogen, hydroxy, protected hydroxy, cyano, nitro, C<sub>1</sub> to C<sub>6</sub> alkyl, C<sub>1</sub> to C<sub>6</sub> substituted alkyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> substituted alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> substituted acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, N, N-di(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, trifluoromethyl, N-((C<sub>1</sub> to C<sub>6</sub> alkyl)sulfonyl)amino, N- (phenylsulfonyl)amino or phenyl, wherein the phenyl is substituted or unsubstituted, such that, for example, a biphenyl results.

35

Examples of the term "substituted phenyl" includes a mono- or di(halo)phenyl group such as 2, 3 or 4-chlorophenyl, 2,6-dichlorophenyl, 2,5-dichlorophenyl, 3,4-dichlorophenyl, 2, 3 or 4-bromophenyl, 3,4-dibromophenyl, 3-chloro-4-fluorophenyl,

5 2, 3 or 4-fluorophenyl and the like; a mono or di(hydroxy)phenyl group such as 2, 3 or  
4-hydroxyphenyl, 2,4-dihydroxyphenyl, the protected-hydroxy derivatives thereof and  
the like; a nitrophenyl group such as 2, 3 or 4-nitrophenyl; a cyanophenyl group, for  
example, 2, 3 or 4-cyanophenyl; a mono- or di(alkyl)phenyl group such as 2, 3 or 4-  
methylphenyl, 2,4-dimethylphenyl, 2, 3 or 4-(iso-propyl)phenyl, 2, 3 or 4-ethylphenyl,  
10 2, 3 or 4-(n-propyl)phenyl and the like; a mono or di(alkoxy)phenyl group, for  
example, 2,6-dimethoxyphenyl, 2, 3 or 4-methoxyphenyl, 2, 3 or 4-ethoxyphenyl, 2, 3  
or 4-(isopropoxy)phenyl, 2, 3 or 4-(t-butoxy)phenyl, 3-ethoxy-4-methoxyphenyl and  
the like; 2, 3 or 4-trifluoromethylphenyl; a mono- or dicarboxyphenyl or (protected  
carboxy)phenyl group such as 2, 3 or 4-carboxyphenyl or 2,4-di(protected  
15 carboxy)phenyl; a mono-or di(hydroxymethyl)phenyl or (protected  
hydroxymethyl)phenyl such as 2, 3, or 4-(protected hydroxymethyl)phenyl or  
3,4-di(hydroxymethyl)phenyl; a mono- or di(aminomethyl)phenyl or (protected  
aminomethyl)phenyl such as 2, 3 or 4-(aminomethyl)phenyl or 2,4-(protected  
aminomethyl)phenyl; or a mono- or di(N-(methylsulfonylamino))phenyl such as 2, 3  
20 or 4-(N-(methylsulfonylamino))phenyl. Also, the term "substituted phenyl" represents  
disubstituted phenyl groups wherein the substituents are different, for example, 3-  
methyl-4-hydroxyphenyl, 3-chloro-4-hydroxyphenyl, 2-methoxy-4-bromophenyl,  
4-ethyl-2-hydroxyphenyl, 3-hydroxy-4-nitrophenyl, 2-hydroxy 4-chlorophenyl and the  
like.

25 The term "heteroaryl" means a heterocyclic aromatic derivative which is a five-  
membered or six-membered ring system having from 1 to 4 heteroatoms, such as  
oxygen, sulfur and/or nitrogen, in particular nitrogen, either alone or in conjunction  
with sulfur or oxygen ring atoms. Examples of heteroaryls include pyridinyl,  
30 pyrimidinyl, and pyrazinyl, pyridazinyl, pyrrolo, furano, thiopheno, oxazolo, isoxazolo,  
phthalimido, thiazolo and the like.

The term "substituted heteroaryl" means the above-described heteroaryl is  
substituted with, for example, one or more, and preferably one or two, substituents  
35 which are the same or different which substituents can be halogen, hydroxy,  
protected hydroxy, cyano, nitro, C<sub>1</sub> to C<sub>6</sub> alkyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> substituted  
alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> substituted acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, carboxy, protected  
carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected



5 hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected  
 (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide,  
 N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, N, N-di(C<sub>1</sub>  
 to C<sub>6</sub> alkyl)carboxamide, trifluoromethyl, N-((C<sub>1</sub> to C<sub>6</sub> alkyl)sulfonyl)amino or N-  
 (phenylsulfonyl)amino groups.

10 The term "substituted naphthyl" specifies a naphthyl group substituted with one or  
 more, and preferably one or two, moieties either on the same ring or on different  
 rings chosen from the groups consisting of halogen, hydroxy, protected hydroxy,  
 cyano, nitro, C<sub>1</sub> to C<sub>6</sub> alkyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, carboxy,  
 15 protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl,  
 protected hydroxymethyl, amino, protected amino, (monosubstituted)amino,  
 protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected  
 carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub>  
 alkyl)carboxamide, N, N-di(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, trifluoromethyl, N-((C<sub>1</sub> to C<sub>6</sub>  
 20 alkyl)sulfonyl)amino or N-(phenylsulfonyl)amino.

Examples of the term "substituted naphthyl" includes a mono or di(halo)naphthyl  
 group such as 1, 2, 3, 4, 5, 6, 7 or 8-chloronaphthyl, 2, 6-dichloronaphthyl, 2, 5-  
 dichloronaphthyl, 3, 4-dichloronaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-bromonaphthyl, 3, 4-  
 25 dibromonaphthyl, 3-chloro-4-fluoronaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-fluoronaphthyl and  
 the like; a mono or di(hydroxy)naphthyl group such as 1, 2, 3, 4, 5, 6, 7 or 8-  
 hydroxynaphthyl, 2, 4-dihydroxynaphthyl, the protected-hydroxy derivatives thereof  
 and the like; a nitronaphthyl group such as 3- or 4-nitronaphthyl; a cyanonaphthyl  
 group, for example, 1, 2, 3, 4, 5, 6, 7 or 8-cyanonaphthyl; a mono- or  
 30 di(alkyl)naphthyl group such as 2, 3, 4, 5, 6, 7 or 8-methylnaphthyl, 1, 2,  
 4-dimethylnaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-(isopropyl)naphthyl, 1, 2, 3, 4, 5, 6, 7 or  
 8-ethylnaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-(n-propyl)naphthyl and the like; a mono or  
 di(alkoxy)naphthyl group, for example, 2, 6-dimethoxynaphthyl, 1, 2, 3, 4, 5, 6, 7 or  
 8-methoxynaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-ethoxynaphthyl, 1, 2, 3, 4, 5, 6, 7 or  
 35 8-(isopropoxy)naphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-(t-butoxy)naphthyl, 3-ethoxy-4-  
 methoxynaphthyl and the like; 1, 2, 3, 4, 5, 6, 7 or 8-trifluoromethylnaphthyl; a mono-  
 or dicarboxynaphthyl or (protected carboxy)naphthyl group such as 1, 2, 3, 4, 5, 6, 7  
 or 8-carboxynaphthyl or 2, 4-di(-protected carboxy)naphthyl; a mono-or

- 5 di(hydroxymethyl)naphthyl or (protected hydroxymethyl)naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(protected hydroxymethyl)naphthyl or 3, 4-di(hydroxymethyl)naphthyl; a mono- or di(amino)naphthyl or (protected amino)naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(amino)naphthyl or 2, 4-(protected amino)-naphthyl, a mono- or di(aminomethyl)naphthyl or (protected aminomethyl)naphthyl such as 2, 3, or 10 4-(aminomethyl)naphthyl or 2, 4-(protected aminomethyl)-naphthyl; or a mono- or di-(N-methylsulfonylamino) naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(N-methylsulfonylamino)naphthyl. Also, the term "substituted naphthyl" represents disubstituted naphthyl groups wherein the substituents are different, for example, 3-methyl-4-hydroxynaphth-1-yl, 3-chloro-4-hydroxynaphth-2-yl, 2-methoxy-4-bromonaphth-1-yl, 4-ethyl-2-hydroxynaphth-1-yl, 3-hydroxy-4-nitronaphth-2-yl, 2-hydroxy-4-chloronaphth-1-yl, 2-methoxy-7-bromonaphth-1-yl, 4-ethyl-5-hydroxynaphth-2-yl, 3-hydroxy-8-nitronaphth-2-yl, 2-hydroxy-5-chloronaphth-1-yl and the like.
- 20 The term "C<sub>1</sub> to C<sub>8</sub> alkyl" denotes such radicals as methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, sec-butyl, tert-butyl, amyl, tert-amyl, hexyl, n-heptyl, 2-heptyl, 3-heptyl, 4-heptyl, 2-methyl-1hexyl, 2-methyl-2hexyl, 2-methyl-3-hexyl, n-octyl and the like.
- 25 The term "C<sub>2</sub> to C<sub>6</sub> alkenyl" denotes such radicals as propenyl or butenyl.

Examples of the above substituted alkyl groups include the 2-oxo-prop-1-yl, 3-oxo-but-1-yl, cyanomethyl, nitromethyl, chloromethyl, hydroxymethyl, tetrahydropyranyloxymethyl, trityloxymethyl, propionyloxymethyl, amino, 30 methylamino, aminomethyl, dimethylamino, carboxymethyl, allyloxycarbonylmethyl, allyloxycarbonylaminomethyl, methoxymethyl, ethoxymethyl, t-butoxymethyl, acetoxymethyl, chloromethyl, bromomethyl, iodomethyl, trifluoromethyl, 6-hydroxyhexyl, 2,4-dichloro(n-butyl), 2-aminopropyl, 1-chloroethyl, 2-chloroethyl, 1-bromoethyl, 2-chloroethyl, 1-fluoroethyl, 2-fluoroethyl, 1-iodoethyl, 2-iodoethyl, 1-chloropropyl, 2-chloropropyl, 3-chloropropyl, 1-bromopropyl, 2-bromopropyl, 3-bromopropyl, 1-fluoropropyl, 2-fluoropropyl, 3-fluoropropyl, 1-iodopropyl, 2-iodopropyl, 3-iodopropyl, 2-aminoethyl, 1-aminoethyl, N-benzoyl-2-aminoethyl, N-acetyl-2-aminoethyl, N-benzoyl-1-aminoethyl, N-acetyl-1-aminoethyl and the like.

5

The term "C<sub>1</sub> to C<sub>8</sub> substituted alkyl" denotes that the above C<sub>1</sub> to C<sub>8</sub> alkyl groups are substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, C<sub>3</sub> to C<sub>7</sub> cycloalkyl, naphthyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, 10 guanidino, protected guanidino, heterocyclic ring, substituted heterocyclic ring, imidazolyl, indolyl, pyrrolidinyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, nitro, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, N,N-di(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C<sub>1</sub> to C<sub>4</sub> alkylthio or C<sub>1</sub> to C<sub>4</sub> 15 alkylsulfonyl groups. The substituted alkyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

The term "C<sub>7</sub> to C<sub>12</sub> phenylalkyl" denotes a C<sub>1</sub> to C<sub>6</sub> alkyl group substituted at any position by a phenyl, substituted phenyl, heteroaryl or substituted heteroaryl.

20 Examples of such a group include benzyl, 2-phenylethyl, 3-phenyl(n-propyl), 4-phenylhexyl, 3-phenyl(n-amyl), 3-phenyl(sec-butyl) and the like. Preferred C<sub>7</sub> to C<sub>12</sub> phenylalkyl groups are the benzyl and the phenylethyl groups.

The term "C<sub>7</sub> to C<sub>12</sub> substituted phenylalkyl" denotes a C<sub>7</sub> to C<sub>12</sub> phenylalkyl group substituted on the C<sub>1</sub> to C<sub>6</sub> alkyl portion with one or more, and preferably one or two, 25 groups chosen from halogen, hydroxy, protected hydroxy, oxo, protected oxo, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, guanidino, protected guanidino, heterocyclic ring, substituted heterocyclic ring, C<sub>1</sub> to C<sub>6</sub> alkyl, C<sub>1</sub> to C<sub>6</sub> substituted alkyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> substituted alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> substituted acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, nitro, 30 carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, N, N-(C<sub>1</sub> to C<sub>6</sub> dialkyl)carboxamide, cyano, N-(C<sub>1</sub> to C<sub>6</sub> alkylsulfonyl)amino, thiol, C<sub>1</sub> to C<sub>4</sub> alkylthio, C<sub>1</sub> to C<sub>4</sub> alkylsulfonyl groups; and/or the phenyl group may be substituted with one or more, and preferably one or two, substituents chosen from halogen, hydroxy, 35 protected hydroxy, cyano, nitro, C<sub>1</sub> to C<sub>6</sub> alkyl, C<sub>1</sub> to C<sub>6</sub> substituted alkyl, C<sub>1</sub> to C<sub>7</sub> alkoxy, C<sub>1</sub> to C<sub>7</sub> substituted alkoxy, C<sub>1</sub> to C<sub>7</sub> acyl, C<sub>1</sub> to C<sub>7</sub> substituted acyl, C<sub>1</sub> to C<sub>7</sub> acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino,

5 (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>6</sub> alkyl) carboxamide, protected N-(C<sub>1</sub> to C<sub>6</sub> alkyl) carboxamide, N, N-di(C<sub>1</sub> to C<sub>6</sub> alkyl)carboxamide, trifluoromethyl, N-((C<sub>1</sub> to C<sub>6</sub> alkyl)sulfonyl)amino, N-(phenylsulfonyl)amino, cyclic C<sub>2</sub> to C<sub>7</sub> alkylene or a phenyl group, substituted or unsubstituted, for a resulting biphenyl group. The  
10 substituted alkyl or phenyl groups may be substituted with one or more, and preferably one or two, substituents which can be the same or different.

Examples of the term "C<sub>7</sub> to C<sub>12</sub> substituted phenylalkyl" include groups such as 2-phenyl-1-chloroethyl, 2-(4-methoxyphenyl)ethyl, 4-(2,6-dihydroxy phenyl)n-hexyl, 2-  
15 (5-cyano-3-methoxyphenyl)n-pentyl, 3-(2,6-dimethylphenyl)n-propyl, 4-chloro-3-aminobenzyl, 6-(4-methoxyphenyl)-3-carboxy(n-hexyl), 5-(4-aminomethylphenyl)- 3-(aminomethyl)n-pentyl, 5-phenyl-3-oxo-n-pent-1-yl and the like.

20 The term "heterocycle" or "heterocyclic ring" denotes optionally substituted five-membered to eight-membered rings that have 1 to 4 heteroatoms, such as oxygen, sulfur and/or nitrogen, in particular nitrogen, either alone or in conjunction with sulfur  
25 or oxygen ring atoms. These five-membered to eight-membered rings may be saturated, fully unsaturated or partially unsaturated, with fully saturated rings being preferred. Preferred heterocyclic rings include morpholino, piperidinyl, piperazinyl, 2-amino-imidazolyl, tetrahydrofurano, pyrrolo, tetrahydrothiophen-yl, hexylmethyleneimino and heptylmethyleneimino.

30 The term "substituted heterocycle" or "substituted heterocyclic ring" means the above-described heterocyclic ring is substituted with, for example, one or more, and preferably one or two, substituents which are the same or different which substituents can be halogen, hydroxy, protected hydroxy, cyano, nitro, C<sub>1</sub> to C<sub>12</sub> alkyl, C<sub>1</sub> to C<sub>12</sub>  
35 alkoxy, C<sub>1</sub> to C<sub>12</sub> substituted alkoxy, C<sub>1</sub> to C<sub>12</sub> acyl, C<sub>1</sub> to C<sub>12</sub> acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino carboxamide, protected

carboxamide, N-(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, N, N-di(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, trifluoromethyl, N-((C<sub>1</sub> to C<sub>12</sub> alkyl)sulfonyl)amino, N-(phenylsulfonyl)amino, heterocycle or substituted heterocycle groups.

The term "C<sub>1</sub> to C<sub>8</sub> alkoxy" as used herein denotes groups such as methoxy, ethoxy, n-propoxy, isopropoxy, n-butoxy, t-butoxy and like groups. A preferred alkoxy is methoxy. The term "C<sub>1</sub> to C<sub>8</sub> substituted alkoxy" means the alkyl portion of the alkoxy can be substituted in the same manner as in relation to C<sub>1</sub> to C<sub>8</sub> substituted alkyl.

The term "C<sub>1</sub> to C<sub>8</sub> aminoacyl" encompasses groups such as formyl, acetyl, propionyl, butyryl, pentanoyl, pivaloyl, hexanoyl, heptanoyl, octanoyl, benzoyl and the like.

The term "C<sub>1</sub> to C<sub>8</sub> substituted aminoacyl" denotes the acyl group substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, cyclohexyl, naphthyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, guanidino, heterocyclic ring, substituted heterocyclic ring, imidazolyl, indolyl, pyrrolidinyl, C<sub>1</sub> to C<sub>12</sub> alkoxy, C<sub>1</sub> to C<sub>12</sub> acyl, C<sub>1</sub> to C<sub>12</sub> acyloxy, nitro, C<sub>1</sub> to C<sub>12</sub> alkyl ester, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, protected N-(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, N,N-di(C<sub>1</sub> to C<sub>12</sub> alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C<sub>1</sub> to C<sub>10</sub> alkylthio or C<sub>1</sub> to C<sub>10</sub> alkylsulfonyl groups. The substituted acyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

Examples of C<sub>1</sub> to C<sub>8</sub> substituted acyl groups include 4-phenylbutyryl, 3-phenylbutyryl, 3-phenylpropanoyl, 2-cyclohexanylacetyl, cyclohexanecarbonyl, 2-furanoyl and 3-dimethylaminobenzoyl.

This invention also provides a pharmaceutical composition comprising an effective amount of a compound according to the invention. Such pharmaceutical compositions can be administered by various routes, for example oral,

5 subcutaneous, intramuscular, intravenous or intracerebral. The preferred route of administration would be oral at daily doses of the compound for adult human treatment of about 0.01-5000 mg, preferably 1-1500 mg per day. The appropriate dose may be administered in a single dose or as divided doses presented at appropriate intervals for example as two, three four or more subdoses per day.

10 For preparing pharmaceutical compositions containing compounds of the invention, inert, pharmaceutically acceptable carriers are used. The pharmaceutical carrier can be either solid or liquid. Solid form preparations include, for example, powders, tablets, dispersible granules, capsules, cachets, and suppositories.

15 A solid carrier can be one or more substances which can also act as diluents, flavoring agents, solubilizers, lubricants, suspending agents, binders, or tablet disintegrating agents; it can also be an encapsulating material.

20 In powders, the carrier is generally a finely divided solid which is in a mixture with the finely divided active component. In tablets, the active compound is mixed with the carrier having the necessary binding properties in suitable proportions and compacted in the shape and size desired.

25 For preparing pharmaceutical composition in the form of suppositories, a low-melting wax such as a mixture of fatty acid glycerides and cocoa butter is first melted and the active ingredient is dispersed therein by, for example, stirring. The molten homogeneous mixture is then poured into convenient-sized molds and allowed to cool and solidify.

30 Powders and tablets preferably contain between about 5% to about 70% by weight of the active ingredient, preferably comprising (especially consisting of) one or more of the compounds according to this invention. Suitable carriers include, for example, magnesium carbonate, magnesium stearate, talc, lactose, sugar, pectin, dextrin,  
35 starch, tragacanth, methyl cellulose, sodium carboxymethyl cellulose, a low-melting wax, cocoa butter and the like.

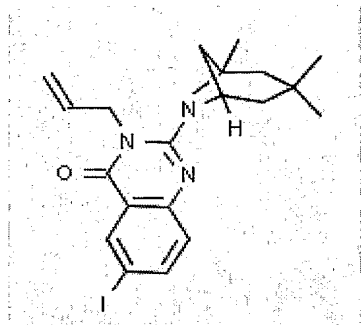
5 The pharmaceutical compositions can include the formulation of the active compound with encapsulating material as a carrier providing a capsule in which the active component (with or without other carriers) is surrounded by a carrier, which is thus in association with it. In a similar manner, cachets are also included. Tablets, powders, cachets, and capsules can be used as solid dosage forms suitable for oral  
10 administration.

Liquid pharmaceutical compositions include, for example, solutions suitable for oral or parenteral administration, or suspensions, and emulsions suitable for oral administration. Sterile water solutions of the active component or sterile solutions of  
15 the active component in solvents comprising water, ethanol, or propylene glycol are examples of liquid compositions suitable for parenteral administration.

Sterile solutions can be prepared by dissolving the active component in the desired solvent system, and then passing the resulting solution through a membrane filter to  
20 sterilize it or, alternatively, by dissolving the sterile compound in a previously sterilized solvent under sterile conditions.

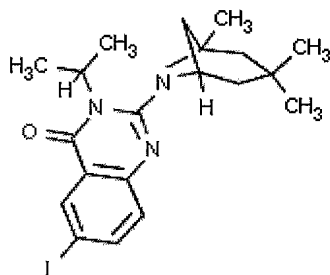
In a preferred embodiment of the invention in the compounds claimed, or the pharmaceutical acceptable salts or solvates thereof,  $R_1$ ,  $R_3$  and  $R_4$  are H,  $R_2$  is  
25 halogen and preferably iodine over bromine and chlorine and  $R_5$  is H,  $C_1$  to  $C_8$  alkyl or  $C_1$  to  $C_8$  substituted alkyl.

A particularly preferred compound which may act as a partial agonist of LXR is shown in formula (II) below (MOLNAME TR1040001892). It has been demonstrated  
30 that this compound has a low effective concentration at LXR with an  $EC_{50}$  of 2  $\mu$ M in a FRET assay wherein the  $EC_{50}$  reflects the half-maximal effective concentration, and which is higher than the  $EC_{50}$  of 0.015  $\mu$ M for the published LXR agonist TO901317 (J. Schultz et al., Genes Dev. 14, 2831-2838, 2000). Compound according to formula (II) does show selective upregulation of ABCA1 and ABCG1 in  
35 THP-1 macrophages but does not significantly upregulate FAS and much reduced SREBP-1c in HepG2 cells (see EXAMPLE 5 further down).



(II)

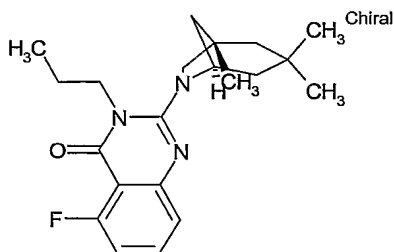
It has also been found that the compound according to formula (III) (Molname TR1040011382) to be active as partial agonist of the LXR human nuclear receptors with a selective upregulation of the target genes ABCA1 and ABCG1 in THP-1 cells compared to FAS and SREBP-1c in HepG2 cells.



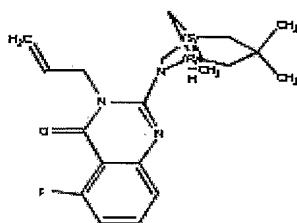
(III)

It has also been found that the compounds according to formulas (IV) (Molname TR1040002211) and (V) (Molname TR1040002212) to be active as partial agonist of the LXR human nuclear receptors however with a reduced selectivity regarding the upregulation of the target genes ABCA1 and ABCG1 in THP-1 cells versus FAS and SREBP-1c in HepG2 cells compared to compounds of formula (II) and (III) (see EXAMPLE 5 further down)





(IV)



(V)

In particular the invention relates to a compound as described above wherein said compounds is capable of binding the LXR receptor protein or a portion thereof encoded by a nucleic acid according to SEQ ID NO:1 or NO:2 (Fig. 3 ) or a mammalian homologue thereof. This compound can bind to the LXR receptor protein or a portion thereof in a mixture comprising 10-200 ng of LXR receptor protein, a fusion protein containing LXR or a portion thereof, preferably the ligand binding domain, fused to a Tag, 5-100 mM Tris/HCl at pH 6,8-8,3; 60-1000 mM KCl; 0-20 mM MgCl<sub>2</sub>; 100-1000ng/μl BSA in a total volume of preferably about 25 μl[see also EXAMPLE 1 and Fig.2).

A mammalian receptor protein homologue of the protein encoded by a nucleic acid according to SEQ ID NO:1 or 2, as used herein is a protein that performs substantially the same task as LXR does in humans and shares at least 40% sequence identity at the amino acid level, preferably over 50 % sequence identity at the amino acid level more preferably over 65 % sequence identity at the amino acid

- 5 level, even more preferably over 75 % sequence identity at the amino acid level and  
most preferably over 85 % sequence identity at the amino acid level.

The invention in particular concerns a method for prevention or treatment of a LXR  
receptor protein or LXR receptor protein homologue mediated disease or condition in  
10 a mammal comprising administration of a therapeutically effective amount of a  
compound according to the invention wherein the prevention or treatment is directly  
or indirectly accomplished through the binding of a compound according to the  
invention to the LXR receptor protein or to the LXR receptor protein homologue.

- 15 The term mediated herein means that the physiological pathway in which the LXR  
receptor protein acts is either directly or indirectly involved in the disease or condition  
to be treated or prevented. In the case where it is indirectly involved it could be that,  
e.g. modulating the activity of LXR by a compound according to the invention  
influences a parameter which has a beneficial effect on a disease or a condition. One  
20 such example is that modulation of LXR activity leads to decreased levels of serum  
cholesterol or certain lipoproteins which in turn have a beneficial effect on the  
prevention and treatment of atherosclerosis. Herein a condition is a physiological or  
phenotypic state which is desirably altered. One such example would be obesity  
which is not necessarily medically harmful but nonetheless a non desirable  
25 phenotypic condition. In a preferred embodiment of the invention the method for  
prevention or treatment of a LXR receptor protein mediated disease or condition is  
applied to a human. This may be male or female.

- Pharmaceutical compositions generally are administered in an amount effective for  
30 treatment or prophylaxis of a specific condition or conditions. Initial dosing in human  
is accompanied by clinical monitoring of symptoms, such symptoms for the selected  
condition. In general, the compositions are administered in an amount of active agent  
of at least about 100 µg/kg body weight. In most cases they will be administered in  
one or more doses in an amount not in excess of about 20 mg/kg body weight per  
35 day. Preferably, in most cases, doses is from about 100 µg/kg to about 5 mg/kg body  
weight, daily.

5 For administration particularly to mammals, and particularly humans, it is expected  
that the daily dosage level of active agent will be 0,1 mg/kg to 10 mg/kg and typically  
around 1 mg/kg.

10 By "therapeutically effective amount" is meant a symptom-alleviating or symptom-  
reducing amount, a cholesterol-reducing amount, a cholesterol absorption blocking  
amount, a protein and/or carbohydrate digestion-blocking amount and/or a de novo  
cholesterol biosynthesis-blocking amount of a compound according to the invention.

15 Likewise the invention concerns a method of treating in mammal a disease which is  
correlated with abnormal cholesterol, triglyceride, or bile acid levels or deposits  
comprising administering to a mammal in need of such treatment a therapeutically  
effective amount of a compound according to the invention.

20 Accordingly, the compounds according to the invention may also be used in a  
method of prevention or treatment of mammalian atherosclerosis, gallstone disease,  
lipid disorders, Alzheimer's disease, skin disorders, inflammation, obesity or  
cardiovascular disorders such as coronary heart disease or stroke.

25 The invention further concerns a method of blocking in a mammal the cholesterol  
absorption in the intestine in need of such blocking comprising administering to a  
mammal in need of such treatment a therapeutically effective amount of a compound  
according to the invention. The invention may also be used to treat obesity in  
humans.

30

The Liver X Receptor alpha is a prototypical type 2 nuclear receptor meaning that it  
activates genes upon binding to the promoter region of target genes in a  
heterodimeric fashion with Retinoid X Receptor. The relevant physiological ligands of  
LXR are oxysterols. The compounds have been demonstrated to have a high binding  
35 efficacy (binding coefficients measured as EC50 in the range of 1-5  $\mu$ M) as well as  
agonistic and/or partial agonistic properties. Consequently they may be applied to  
regulate genes that participate in bile acid, cholesterol and fatty acid homeostasis as  
well as other downstream regulated genes. Examples of such genes are but are not

5 limited to lipid absorption, cholesterol biosynthesis, cholesterol transport or binding,  
bile acid transport or binding, proteolysis, amino acid metabolism, glucose  
biosynthesis, protein translation, electron transport, and hepatic fatty acid  
metabolism. LXR often functions in vivo as a heterodimer with the Retinoid X  
Receptor. Published non-steroidal LXR agonists such as the "Tularik" compound  
10 "TO901317" (see figure 5) are known to influence the regulation of various liver  
genes. Genes found to be regulated by TO901317 can be found in figure 6. Thus, the  
invention also concerns a method of modulating a gene whose expression is  
regulated by the LXR receptor in a mammal comprising administration of a  
therapeutically effective amount of a compound according to the invention to said  
15 mammal.

A number of direct and indirect LXR target genes have been described whose  
regulated expression contribute to cholesterol homeostasis and lipogenesis. In this  
respect the direct regulation of Cyp7A, which was shown to be a direct target gene of  
20 LXR at least in the rodent lineage is an important aspect of cholesterol removal by  
increased metabolism of bile acids (Lehmann et al., J Biol.Chem. 272 (6) 3137-3140;  
1007). Gupta et al. (Biochem. Biophys Res.Com, 293; 338-343, 2002) showed that  
LXR  $\alpha$  regulation of Cyp7A is dominant over FXR inhibitory effects on Cyp7A  
transcription.

25 A key transcription factor that was also shown to be a direct target gene for the LXR  
receptor is SREBP-1C (Repa et al., Genes and Development, 14:2819-2830; 2000;  
Yoshikawa et al.; Mol.Cell.Biol.21 (9) 2991-3000, 2001). SREBP-1C itself activates  
transcription of genes involved in cholesterol and fatty acid synthesis in liver but also  
30 other mammalian tissues. Some of the SREBP1c target genes involved in  
lipogenesis like FAS and SCD have shown to be additionally direct targets of the  
LXR receptors (Joseph et al.; J Biol Chem. 2002 Mar 29;277(13):11019-25; Liang et  
al., J Biol Chem. 2002 Mar 15;277(11):9520-8.).

35 A primary limitation for the applicability of LXR agonists as e.g. anti-atherosclerotic  
drugs comes from the observation that compounds with full agonistic activity, e.g.  
TO901317, not only elevate HDL cholesterol levels but do also increase plasma  
triglyceride levels in mice (Schultz et al., 2000 Genes Dev. 14:2831-8.).  
Concommitantly, not only genes that are involved in cholesterol efflux such as the

cholesterol transporter ABCA1 (Venkateswaran et al., 2000 PNAS. 97:12097-102.), ABCG1 as well as the lipid binding protein Apolipoprotein E (Laffite et al. 2001 PNAS 98:507-512) are induced by full LXR agonists, but also genes involved in lipogenesis, including the fatty acid synthase FAS (Joseph et al 2002 J Biol Chem. 277:11019-11025), and SREB P-1c (Yoshikawa et al., 2001 Mol Cell Biol. 21:2991-3000). Elevation of serum triglyceride levels is an independent risk factor for atherosclerosis (for review see Miller (1999) Hosp Pract (Off Ed) 34: 67-73.). Thus, LXR activity needs to be selectively modulated for therapeutic benefit. In particular, compounds need to be found that stimulate reverse cholesterol transport, but do not significantly increase triglyceride levels.

Another gene that has been shown to be directly regulated by LXRs is the LPL gene, that codes for a key enzyme that is responsible for the hydrolysis of triglycerides in circulating lipoprotein, releasing free fatty acids to peripheral tissues. (Zhang et al. J Biol Chem. 2001 Nov 16;276(46):43018-24.) This enzyme is believed to promote uptake of HDL cholesterol in liver, thereby promoting reverse cholesterol transport. A similar functional involvement in HDL clearance is described for the CETP gene product that facilitated the transfer of HDL cholesterol esters from plasma to the liver. LXR response elements were found in the CETP promoter and direct activation of this gene by LXR was demonstrated (Luo and Tall; J Clin Invest. 2000 Feb;105(4):513-20.).

The regulated transport of cholesterol through biological membranes is an important mechanism in order to maintain cholesterol homeostasis. A pivotal role in these processes in multiple tissues like e.g. macrophages and intestinal mucosa cells is maintained by the ATP-binding cassette transporter proteins (ABC). ABCA1 and ABCG1 were identified as direct LXR target genes (Costet et al.; J Biol Chem. 2000 Sep 8;275(36):28240-5) that mediate cholesterol efflux and prevent thereby e.g. generation of atherogenic plaques in macrophages (Singaraja et al. J Clin Invest. 2002 Jul;110(1):35-42). Other ABC transporters like ABCG5 and ABCG8, primarily expressed in hepatocytes and enterocytes have also been reported to be directly responsive to LXR agonists (Repa et al., J Biol Chem. 2002 May 24;277(21):18793-800. Kennedy et al., J Biol Chem. 2001 Oct 19;276(42):39438-47) and mediate the secretion of sterols from the liver and efflux of dietary sterols from the gut.

5 Apolipoproteins E, C-I, C-II, and C-IV, that fulfill important roles in lipoprotein/lipid  
homeostasis have also been shown to be direct targets of the LXR receptor ( Laffitte  
et al., Proc Natl Acad Sci U S A. 2001 Jan 16;98(2):507-12; Mak et al.; J Biol Chem.  
2002 May 24 [epub ahead of print]). These proteins have been found to be crucial  
10 components of chylomicrons, VLDL, IDL, and HDL and are among other things  
associated with hypertriglyceridemia and arteriosclerosis.

Recently the LXR $\alpha$  itself was shown to be regulated by both LXR receptors in human  
cell types including macrophages suggesting an autoregulatory amplification event in  
the response to LXR ligands which could e.g. lead to an enhanced stimulation of LXR  
15 target genes like e.g. ABCA1 (Bolten et al.; Mol Endocrinol. 2002 Mar;16(3):506-14.;  
Laffitte et al., Mol Cell Biol. 2001 Nov;21(22):7558-68; Whitney et al.; J Biol Chem.  
2001 Nov 23;276(47):43509-15).

Besides the important function of LXR receptors in tissues like liver and  
20 macrophages it has recently been reported that that stimulation of epidermal  
differentiation is mediated by Liver X receptors in murine epidermis. Differentiation  
marker genes like involucrin, loricin and profilaggrin have been shown to be  
upregulated upon LXR ligand treatment (Kömüves et al.; J Invest Dermatol. 2002  
Jan;118(1):25-34.).

25 Another recent report describes the regulation of cholesterol homeostasis (primarily  
the regulation of ABCA1, ABCG1 and SREBP-1C) by the LXR receptors in the  
central nervous system suggesting that LXRs may prove beneficial in the treatment of  
CNS diseases such as Alzheimer's and Niemann-Pick disease that are known to be  
30 accompanied by dysregulation of cholesterol balance (Whitney et al.; Mol Endocrinol.  
2002 Jun;16(6):1378-85).

Activation of LXR by an agonist improves glucose tolerance in a murine model of  
diet-induced obesity and insulin resistance. Gene expression analysis in LXR  
35 agonist-treated mice reveals coordinate regulation of genes involved in glucose  
metabolism in liver and adipose tissue, e.g. the down-regulation of peroxisome  
proliferator-activated receptor gamma coactivator-1 alpha (PGC-1),  
phosphoenolpyruvate carboxykinase (PEPCK), and glucose-6-phosphatase

5 expression and induction of glucokinase in liver. In adipose tissue, activation of LXR led to the transcriptional induction of the insulin-sensitive glucose transporter, GLUT4. LXR agonist may limit hepatic glucose output and improve peripheral glucose uptake (Laffitte et al. (2003) PNAS 100:5419-24).

10 Therefore one other important embodiment of the invention concerns methods that enhances or suppresses amongst other today yet unknown LXR target genes the above mentioned genes and the associated biological processes and pathways through LXR compounds that are subject of this invention.

15 The compounds according to the invention may be used as medicaments, in particular for the manufacture of a medicament for the prevention or treatment of a LXR receptor protein or LXR receptor protein homologue mediated disease or condition in a mammal wherein the prevention or treatment is directly or indirectly accomplished through the binding of the compound according to the invention to the LXR receptor protein or LXR receptor protein homologue. These pharmaceutical  
20 compositions contain 0,1 % to 99,5 % of the compound according to the invention, more particularly 0,5 % to 90 % of the compound according to the invention in combination with a pharmaceutically acceptable carrier.

25 The invention concerns also the use of a compound according to the invention for the manufacture of a medicament for the prevention or treatment of a LXR receptor protein mediated disease or condition wherein the mammal described above is a human. The medicament may be used for regulating the cholesterol transport system, for regulating levels of cholesterol, triglyceride, and/or bile acid in a mammal preferentially a human by activating the LXR receptor. The medicament may be used  
30 for the treatment of atherosclerosis, gallstone disease, lipid disorders, Alzheimer's disease, skin disorders, obesity or a cardiovascular disorder.

35 The invention further concerns the use of a compound according to the invention for the manufacture of a medicament capable for blocking in a mammal, preferentially in a human the cholesterol absorption in the intestine. Further the claimed compound may be used for the manufacture of a medicament for treating obesity in humans and

- 5 for modulating a gene whose expression is regulated by the LXR receptor (see details above and figures).



## EXAMPLES

### EXAMPLE 1

In vitro screening for compounds which influence LXR binding to coactivators.

For screening purposes a GST and 6 x His fusion of the LBD (from amino acids 155 of hLXRalpha to 447) of human LXRalpha is constructed by first cloning a Gateway cassette (Invitrogen) in frame into the Sma I site of the pAGGHLT Polylinker (Pharmingen). Then a PCR fragment specifically amplified from human liver cDNA is cloned into the resulting pACGHLT-GW following the manufacturers instructions for Gateway cloning (Invitrogen) to yield pACGHLT-GW-hLXRalphaLBD.

Primers used for Amplification are:

GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCTTCGCAAATGCCGTCAG and  
GGGGACCACTTTGTACAAGAAAGCTGGGTCCCCTTCTCAGTCTGTTCCACTT .

100 % sequence integrity of all recombinant products is verified by sequencing.

Recombinant Baculovirus is constructed from pACGHLT-GW-hLXRalphaLBD using the Pharmingen Baculovirus Expression vector system according to instructions of the manufacturer. Monolayer cultures of SF9 cells are infected by the virus as recommended by Pharmingen or 200ml cultures of  $1 \times 10^6$  cells/ml grown in 2 liter Erlenmeyer flasks on an orbital shaker at 30 rpm are infected by 10ml of same virus stock. In both cases cells are harvested 3 days after infection. All cell growth is performed in Gibco SF900 II with Glutamine (Invitrogen) medium without serum supplementation at 28°C. Since SF9 cells contain significant amounts of endogenous GST, purification is performed via His and not via GST affinity chromatography. To this end instructions of Pharmingen for purification of recombinant His tagged proteins from SF9 cells are followed with the following modifications: All detergents are omitted from the buffers and cells were lysed on ice by 5 subsequent sonication pulses using a sonicator needle at maximum power.

All eluates are dialyzed against 20 mM Tris/HCl pH 6,8, 300 mM KCl; 5 mM  $MgCl_2$ ; 1 mM DTT; 0,2 mM PMSF; 10% Glycerol. A typical dialyzed eluate fraction contains the fusion protein at a purity of more than 80%. Total protein concentration is 0,1-0,3 mg/ml.

For E. coli expression of a NR coactivator, pDest17-hTif2BD expressing a NR interaction domain from amino acids 548-878 of human Tif2 (Acc. No: XM\_011633 RefSeq) tagged by 6 N-terminal His residues is constructed. Therefore, a PCR fragment specifically amplified from human liver cDNA is subcloned into pDest 17 (Invitrogen) following the manufacturers instructions for Gateway cloning (Invitrogen). Primers used for Amplification are:  
GGGGACAAGTTTGTACAAAAAAGCAGGCTCGTTAGGGTCATCGTTGGCTTCACC  
and  
GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAAGTTGCCCTGGTCGTGGGT  
TA

For E. coli expression plasmid DNA is transformed into chemically competent E. coli BL21 (Invitrogen, USA) and cells are grown to an OD600 of 0.4-0.7 before expression was induced by addition of 0,5 mM IPTG according instructions of the manufacturer (Invitrogen). After induction for 8 hours at 30°C cells are harvested by centrifugation for 10 minutes at 5000 x g. Fusion proteins are affinity purified using Ni-NTA Agarose (QIAGEN) according to the instructions of the manufacturer. Recombinant Tif2 construct is dialyzed against 20 mM Tris/HCL pH 7.9; 60 mM KCl; 5 mM MgCl<sub>2</sub>; 1 mM DTT, 0,2 mM PMSF; 10% glycerol. A typical dialyzed eluate fraction contains the fusion protein at a purity of more than 80%. Total protein concentration is 0,1-0,3 mg/ml.

The TIF2 fragment is subsequently biotinylated by addition of 5-40µl/ml Tif2 fraction of a Biotinamidocaproate N-Hydroxysuccinimide-ester (Sigma) solution (20 mg/ml in DMSO). Overhead rotating samples are incubated for 2 hours at room temperature. Unincorporated label is then separated using G25 Gel filtration chromatography (Pharmacia Biotech, Sweden). Protein containing fractions from the column are pooled and tested for activity in the assay as described below.

For screening of compound libraries as provided for by the methods shown below in the examples for substances which influence the LXR/Tif 2 interaction, the Perkin Elmer LANCE technology is applied. This method relies on the binding dependent energy transfer from a donor to an acceptor fluorophore attached to the binding partners of interest. For ease of handling and reduction of background from compound fluorescence LANCE technology makes use of generic fluorophore labels and time resolved detection (for detailed description see Hemmilä I, Blomberg K and

Hurskainen P, Time-resolved resonance energy transfer (TR-FRET) principle in LANCE, Abstract of Papers Presented at the 3<sup>rd</sup> Annual Conference of the Society for Biomolecular Screening, Sep., California (1997) )

For screening, 20-200 ng of biotinylated Tif 2 fragment and 10-200 ng of GST-LXR fragment are combined with 0.5-2 nM LANCE Eu-(W1024) labelled anti-GST antibody (Perkin Elmer) and 0,1-0,5µg of highly fluorescent APC-labelled streptavidin (Perkin Elmer, AD0059) in the presence of 50µM of individual compounds to be screened in a total volume of 25 µl of 20 mM Tris /HCl pH 6,8; 300 mM KCl; 5 mM MgCl<sub>2</sub>; 100-1000 ng/µl/ BSA DMSO content of the samples is kept below 4%. Samples are incubated for a minimum of 60 minutes in the dark at room temperature in FIA-Plates black 384well med. binding (Greiner).

The LANCE signal is detected by a Perkin Elmer VICTOR2V™ Multilabel Counter applying the detection parameters listed in Fig. 2. The results are visualized by plotting the ratio between the emitted light at 665 nm and at 615 nm. For every batch of recombinant proteins amount of proteins, including BSA and labeling reagents giving the most sensitive detection of hits is determined individually by analysis of dose response curves for 22R Hydroxycholesterol and TO 901317

## EXAMPLE 2

Experimental procedure for the preparation of the compounds according to the invention.

### *o*-AZIDOBENZOIC ACID SYNTHESIS (2)

The anthranilic acid (1, 1 eq., 0.5-1 M) is suspended in 6 M HCl, containing enough AcOH (0-20% dependent upon the anthranilic acid) to facilitate dissolution of the anthranilic acid and/or the intermediate diazonium salt, and cooled to 0 °C. NaNO<sub>2</sub> (1.1 eq., 1.3-2.5 M) dissolved in H<sub>2</sub>O is added to the anthranilic acid solution at a rate such that the temperature of the reaction solution remains below 5 °C. The resulting homogeneous solution of the diazonium salt is slowly filtered through a sintered glass funnel into a solution of NaN<sub>3</sub> (1.1 eq., 0.7-1.1 M) and NaOAc (12 eq.) in H<sub>2</sub>O. The

reaction mixture is stirred/shaken for 30-60 min following cessation of vigorous N<sub>2</sub> evolution. Following acidification of the reaction mixture to pH 1 with concentrated HCl, the mixture was cooled to 0 °C to encourage complete precipitation of the *o*-azidobenzoic acid. The precipitate is collected by filtration and washed with 6 M HCl (2x) and H<sub>2</sub>O (2x). The *o*-azidobenzoic acid product (2) is dried *in vacuo* (500 mtorr, 30 °C).

#### ACYLATION OF HYDROXYMETHYL RESIN (4)

To hydroxymethyl resin (1.0 eq., 1.3 mmol/g) and the *o*-azidobenzoic acid (1, 2.5 eq.) is added DMF (to give 400 mM *o*-azidobenzoic acid, 1), CsCO<sub>3</sub> (2.0 eq.) and KI (2.0 eq.). Following agitation of the reaction mixture for 36-48 h, the resin-bound *o*-azidobenzoic acid (4) is washed with MeOH (2 cycles), CH<sub>2</sub>Cl<sub>2</sub> (3 cycles), MeOH (3 cycles), DMF (3 cycles), MeOH (3 cycles) and CH<sub>2</sub>Cl<sub>2</sub> (3 cycles), and dried *in vacuo*.

#### AZA-WITTIG FORMATION (5)

To the resin-bound *o*-azidobenzoic acid (4, 1.0 eq.) is added a solution of PPh<sub>3</sub> (THF, 500 mM, 5.0 eq.). After 6 h, the resin is washed with 3 cycles of the following: THF (3 cycles), toluene (3 cycles), CH<sub>2</sub>Cl<sub>2</sub> (3 cycles) and hexanes (3 cycles). Followed by drying *in vacuo* to afford resin bound iminophosphorane (5)

#### CARBODIIMIDE FORMATION (6)

To the resin-bound iminophosphorane (5, 1 eq.) is added isocyanate (9, 5 eq., 450 mM) dissolved in ClCH<sub>2</sub>CH<sub>2</sub>Cl. The compounds are shaken at ambient temperature for 16 h, washed with 3 cycles of the following: THF (3 cycles), toluene (3 cycles), CH<sub>2</sub>Cl<sub>2</sub> (3 cycles) and hexanes (3 cycles), and dried *in vacuo* to afford carbodiimide (6).

#### GUANIDINE FORMATION / CYCLIZATION

To the carbodiimide functionalized resin (6) is added secondary amine (10, 0.6 eq., 500 mM) dissolved in  $\text{ClCH}_2\text{CH}_2\text{Cl}$ . The reaction mixture is heated to 50 °C in an incubator for 12-72 h to afford 2-aminoquinazoline (8).

All of the final products are analyzed using an Evaporative Light Scattering Detector (ELSD) detection to determine purity.

### EXAMPLE 3

This example illustrates that compounds according to the invention (experiments shown were done with MOLSTRUCTURE TR1040001892, TR1040011382, TR1040002211 and TR1040002212 (see formulas (II) to (V) for structural formulas)) activate luciferase reporter gene expression in a dose dependent manner mediated through GAL4-LXR $\alpha$ -LBD or GAL4-LXR $\beta$ -LBD constructs in HEK293 cells. TR1040001892 and TR1040011382 do activate LXR beta LBD truct mediated luciferase activity much stronger than with LXR alpha construct which is in contrast to the similar activation of both LXR alpha and LXR beta LBD containing constructs by TR1040002211 and TO901317.

HEK293 cells are grown in 96 well plates and co-transfected with pFR-luc (Stratagene) and pCMV-BD-LXR $\alpha$ -LBD or pCMV-BD-LXR $\beta$ -LBD (each 100 ng of plasmid DNA per well). Transfection is carried using Lipfectamine 2000 (Gibco-BRL) according to the manufacturers protocol. The ligand binding domains (LBD) of LXR $\alpha$  and LXR $\beta$  are cloned into the pCMV-BD-GW (the Gateway Reading Frame Cassette B is cloned as an EcoRV fragment into SmaI site of pCMV-BD) applying the manufacturer protocols for the Gateway<sup>TM</sup> system (Invitrogen).

Luciferase reporter activity is measured in triplicates from extracts of cells after incubating cells in culture medium (DMEM [Gibco-BRL] + 10% FCS [PAA laboratories]) for 16 hours (5%  $\text{CO}_2$ , 37°C) containing 0,5% DMSO (control) or 0,5% DMSO with increasing concentrations of TR1040001892, TR1040002211 or TO901317 (Sigma T 2320, see figure 5 for structural formula). The type of assay used here is a mammalian one hybrid (M1H) assay that is known to those skilled in the art.

Dose-dependent luciferase activities originating from pFR-luc demonstrate the relative activities of the compounds with the LXRA or LXRB LBDs in this mammalian one hybrid type approach.

#### EXAMPLE 4

This example shows that described compounds can increase the abundance of mRNA of LXR target genes like ABCA1 and ABCG1 in THP-1 cells which are treated with TPA or FAS and SREBP-1c in HepG2 cells as shown in Fig 8A-D.

THP-1 cells are seeded in 24 well plates at  $3 \times 10^5$  cells per well in RPMI 1640 medium containing 10 % FCS and 100 nM TPA for 24 h. HepG2 cells are seeded in poly-L-Lysine coated 24well plates at  $1 \times 10^6$  cells per well in EMEM medium containing 10 % FCS until they are appr. 60% confluent.

Before treatment with LXR compounds, the growth medium is changed to medium containing 10% charcoal/dextran-stripped FCS for 12 h. Treatment is done for 12h (THP-1 cells) and 24h (HepG2 cells), respectively, in medium containing 10% charcoal/dextran-stripped FCS (and 100 nM PMA in the case of THP-1 cells).

LXR compounds are dissolved in DMSO, with the final solvent concentration never exceeding 0.125% . All treatments are done in triplicates and experiments repeated twice. Total RNA is extracted using the Qiagen Rneasy Mini Kit and treated with DNase (DNAfree kit, Ambion). RNA is reverse transcribed with Oligo(dT) primer and real-time reverse transcription PCR (TaqMan) is performed using the ABI Prism 7900HT Sequence Detection System and reagents supplied by Applied Biosystems. mRNA steady state levels are normalised to H3 histone (H3F3A ) expression levels.

The sequences of forward primers, reverse primers and TaqMan probes are as follows :

FAS : CTGAGACGGAGGCCATATGCT, GCTGCCACACGCTCCTCTAG, FAM-CAGCAGTTCACGGACATGGAGCACAA-TAMRA

ABCA1 :TCCTGTGGTGTTTCTGGATGAAC, CTTGACAACACTTAGGGCACAATTC, FAM- ACCACAGGCATGGATCCCAAAGCC-TAMRA

## EXAMPLE 5

All compounds T0901317, TR1040002211, TR1040002212, TR1040001892, TR1040011382, LN0000006662 and LN0000006674 cause a marked increase in cholesterol export in differentiated THP-1 macrophages (see Figure 9 A).

Strikingly, the compound T0901317 causes a marked increase in triglyceride mass in HepG2 liver cells, while compounds like TR1040001892 and TR1040011382 do not cause a significant increase in triglyceride mass. Compounds TR1040002211 and TR1040002212 cause a slight increase in triglyceride mass.

This behavior is similar to the selective transcriptional effect of compounds like TR1040001892 and TR1040011382 on the LXR target genes in HepG2 versus THP-1 cells.

Methods: Cultures of the monocyte-macrophage cell line and the hepatocytes HepG2 are obtained from the American Type tissue Culture Collection, Rockville, MD and were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum, 10 mM HEPES, 2 mM Pyruvate, 50  $\mu$ M  $\beta$ -Mercaptoethanol (THP-1) and Minimum essential medium (Eagle) with 2 mM L-glutamine and Earle's BSS supplemented with 10% fetal bovine serum, 2 mM glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate (HepG2), respectively, at 37°C in 5% CO<sub>2</sub>.

THP-1 cells are differentiated into macrophages by addition of 100 nM Phorbol 12-Myristate 13-Acetate (PMA; Sigma P8139) and PMA included in the medium of all subsequent experiments to maintain differentiation.

For cholesterol efflux measurements and triglyceride analysis cells are seeded in 6well plates at  $1.8 \times 10^6$  cells (THP-1) and  $1 \times 10^6$  cells (HepG2) per well, respectively.

## CHOLESTEROL EFFLUX

THP-1 cells are seeded in 6well plates at  $1 \times 10^6$  cells per well in RPMI 1640 medium containing 10 % FCS and 100 nM TPA for 72 h. After washing with PBS cells are incubated 24 h with fresh in RPMI 1640 medium containing 10 % FCS and 100 nM TPA. Cells are washed twice with PBS and RPMI 1640 medium containing 0.15% BSA and 100 nM TPA is added for further 24 h. Treatment with LXR compounds is done for 24 h in RPMI 1640 medium containing 0.15% BSA, 100 nM TPA and 40

µg/ml ApoA1 (Calbiochem # 178452). Medium is collected, centrifuged to remove cell debris and assayed for cholesterol using a commercial fluorometric kit (Molecular Probes A-12216). The remaining cellular proteins are lysed with 0.3 N NaOH and protein content measured with the Biorad Bradford reagent.

#### TRIGLYCERIDE ASSAY

HepG2 cells are seeded in poly-L-Lysine coated 6well plates at  $1 \times 10^6$  cells per well in EMEM medium containing 10 % FCS until they were appr. 60% confluent. Before treatment with LXR compounds growth medium is changed to medium containing 10% charcoal/dextran-stripped FCS for 12 h. Treatment is done for 24h in medium containing 10% charcoal/dextran-stripped FCS. Cells are washed twice with ice-cold PBS/0.2% BSA and twice with cold PBS and all liquid carefully removed. Triglyceride are extracted with 1.5 ml hexane/ isopropanol = 3:2 per well with gentle shaking for 2-3 h at RT according to Pan et al. (2002) JBC 277, 4413-4421 and Goti et al. (1998) Biochem J, 332 , 57-65 . The extraction solution is collected, dried under vacuum and redissolved in isopropanol/ 1.5% triton. The remaining cellular proteins are lysed with 0.3 N NaOH and protein content measured with the Biorad Bradford reagent.

Triglyceride levels are measured as esterified glycerol using a commercial enzymatic colorimetric kit (Sigma 343-25P). In a preliminary assay it is checked by omitting the lipase enzyme that contribution of free glycerol is negligible.

#### FIGURE CAPTIONS

##### FIG. 1

Fig. 1 shows the synthesis of the compounds according to the invention as also described in EXAMPLE 2.



**FIG. 2**

Fig. 2 shows the measurement parameters employed by the Wallace VICTOR2V™ Multilabel Counter which was used for measuring the EC<sub>50</sub> values (see also EXAMPLE 1)

**FIG. 3**

Shows a table with the accession numbers for the key genes

**FIG. 4**

Fig. 4 shows the internal molecular name used by the applicant (MOLNAME) as well as the corresponding structures of preferred compounds according to the invention. The figure further shows their respective EC<sub>50</sub> values (EC50 AVG) as established according to the Example 1 in multiple experiments (see above), as well as their respective average efficacy (% activity relative to TO901317 control agonist).

**FIG. 5**

Figure 5 shows various known LXR ligands. The compound TO901317 is used as a reference compound here. It is apparent from their structures that the inventors have identified novel compounds which are structurally not related to these known ligands.

**FIG. 6**

Figure 6 shows various genes that have been found to be regulated through binding of an LXR agonist to the LXR protein.

**FIG. 7**

Fig 7A and Fig 7B show dose dependence of indicated compounds with LXR alpha LBD (7A) or LXR beta (7B) LBD containing constructs in mammalian one hybrid (M1H) type assays. The respective  $\mu$ M concentrations of the compounds TO901317,

TR1040002211, TR1040002212. TR1040001892 and TR1040011382 are given on the x-axis and the relative light units (RLU) are depicted on the y-axis.

**Fig. 8**

Analysis of mRNA content of the indicated genes (ABCA1, ABCG1, FAS and SREBP-1c) in total RNA isolated from THP-1 cells (8A and 8B) or HepG2 cells (8C and 8D) treated for 12 or 24 hours with indicated concentrations ( $\mu$ M on x-axis) of T0901317, TR1040011382, TR1040001892 and TR1040002211. The relative fold inuction is depicted on the y-axis.

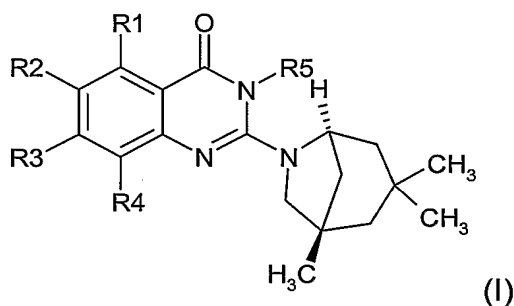
**Fig. 9**

Analysis of relative fold increase in total cholesterol from supernatants of cultivated THP-1 cells (indicated on the y-axis) incubated with ApoA1 and with or without 10 $\mu$ M of the compounds T0901317, TR1040002211, TR1040002212. TR1040001892, TR1040011382, LN0000006662 and LN0000006674 as indicated on the X-axis of Fig 9A.

Analysis of relative levels of total triglyceride (TG) content in HepG2 cells (indicated on the y-axis) treated with 25 $\mu$ M of the indicated compounds T0901317, TR1040002211, TR1040002212. TR1040001892, TR1040011382, LN0000006662 and LN0000006674 (indicated on the x-axis).

**Claims:**

1. A compound of the formula (I), or pharmaceutically acceptable salts or solvates thereof,

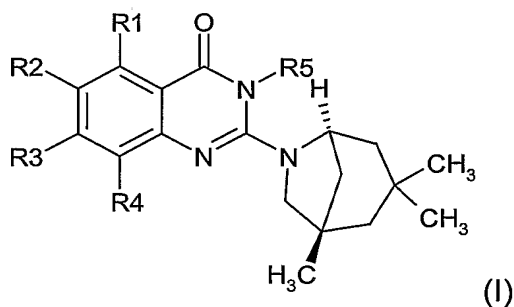


wherein substituents and indices have the following meanings:

$R_1$ ,  $R_2$ ,  $R_3$ ,  $R_4$ , - independent from each other - is H, halogen, hydroxy, protected hydroxy, cyano, nitro,  $C_1$  to  $C_6$  alkyl,  $C_1$  to  $C_6$  substituted alkyl,  $C_1$  to  $C_7$  alkoxy,  $C_1$  to  $C_7$  substituted alkoxy,  $C_1$  to  $C_7$  acyl,  $C_1$  to  $C_7$  substituted acyl,  $C_1$  to  $C_7$  acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-( $C_1$  to  $C_6$  alkyl)carboxamide, protected N-( $C_1$  to  $C_6$  alkyl)carboxamide, N, N-di( $C_1$  to  $C_6$  alkyl)carboxamide, trifluoromethyl, N-(( $C_1$  to  $C_6$  alkyl)sulfonyl)amino, N-(phenylsulfonyl)amino or phenyl,

$R_5$  is H,  $C_1$  to  $C_8$  alkyl,  $C_2$  to  $C_6$  alkenyl,  $C_1$  to  $C_8$  substituted alkyl,  $C_7$  to  $C_{12}$  alkylphenyl or  $C_7$  to  $C_{12}$  substituted phenylalkyl.

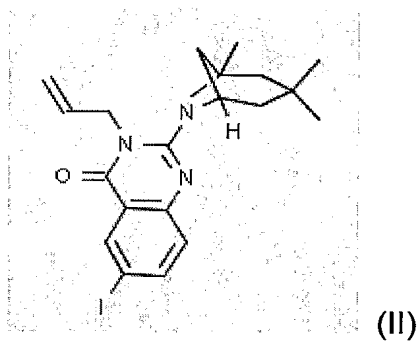
2. A compound of the formula (I), or pharmaceutically acceptable salts or solvates thereof,



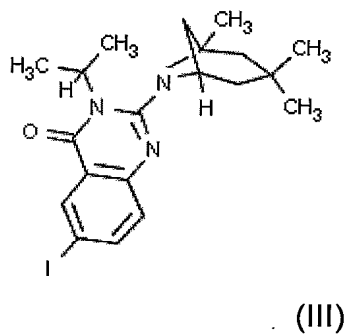
wherein the substituents and indices have the following meanings:

$R_1$ ,  $R_3$  and  $R_4$  are H,  $R_2$  is halogen, hydroxy, protected hydroxy, cyano, nitro,  $C_1$  to  $C_6$  alkyl,  $C_1$  to  $C_6$  substituted alkyl,  $C_1$  to  $C_7$  alkoxy,  $C_1$  to  $C_7$  substituted alkoxy,  $C_1$  to  $C_7$  acyl,  $C_1$  to  $C_7$  substituted acyl,  $C_1$  to  $C_7$  acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-( $C_1$  to  $C_6$  alkyl)carboxamide, protected N-( $C_1$  to  $C_6$  alkyl)carboxamide, N, N-di( $C_1$  to  $C_6$  alkyl)carboxamide, trifluoromethyl, N-(( $C_1$  to  $C_6$  alkyl)sulfonyl)amino, N- (phenylsulfonyl)amino or phenyl, and  $R_5$  is H,  $C_1$  to  $C_8$  alkyl,  $C_2$  to  $C_6$  alkenyl,  $C_1$  to  $C_8$  substituted alkyl.

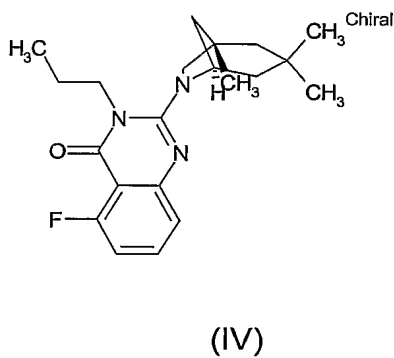
3. A compound according to claim 1 or 2 with:  $R_1$ ,  $R_3$  and  $R_4$  being H,  $R_2$  being halogen and preferably iodine over bromine and chlorine and  $R_5$  being H,  $C_1$  to  $C_6$  alkyl,  $C_3$  to  $C_5$  alkenyl,  $C_1$  to  $C_6$  substituted alkyl.
4. A compound according to any of claims 1 to 3 being



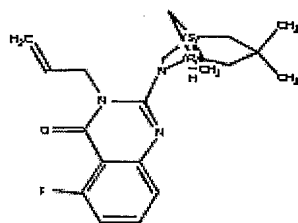
5. A compound according to any of claims 1 to 3 being



6. A compound according to claim 1 being

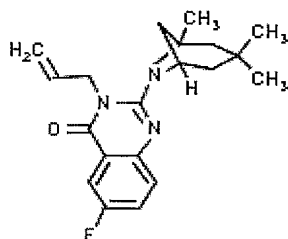


7. A compound according to claim 1 being



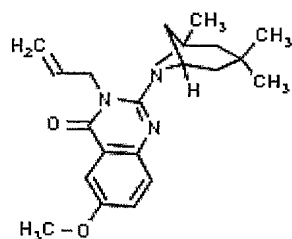
(V)

8. A compound according to any of claims 1 to 3 being



(VI)

9. A compound according to claim 1 or 2 being



(VII)

10. A compound of according to any of claims 1 to 9 wherein said compound is capable of binding the NR1H3 receptor protein encoded by a nucleic acid comprising SEQ ID NO:2 or a portion thereof or a mammalian homologue thereof .
11. A compound of according to any of claims 1 to 9 wherein said compound is capable of binding the NR1H2 receptor protein encoded by a nucleic acid comprising SEQ ID NO:1 or a portion thereof or a mammalian homologue thereof.
12. A compound according to any of claims 1 to 9 for use as a medicament.
13. A method for prevention or treatment of a NR1H3 and/or NR1H2 receptor protein mediated disease or NR1H3 and/or NR1H2 receptor protein homologue mediated disease or condition in a mammal comprising administration of a therapeutically effective amount of a compound according to claims 1 to 9 wherein the prevention or treatment is directly or indirectly accomplished through the binding of the compound according claims 1 to 9 to the NR1H3 and/or NR1H2 receptor proteins or to the NR1H3 and/or NR1H2 receptor protein homologues.
14. A method for regulating the cholesterol synthesis or transport in a mammal which comprises activating the NR1H3 and/or NR1H2 receptors with a therapeutically effective amount of a compound according to claims 1 to 9.
15. A method of treating in mammal a disease which is affected by cholesterol, triglyceride, bile acid, glucose or glucocorticoid levels comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound according to claims 1 to 9.
16. A method of treating in a mammal Atherosclerosis, Alzheimers disease, Type II diabetes, lipid disorders, obesity, an inflammatory or a cardiovascular

disorder comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound according to claims 1 to 9.

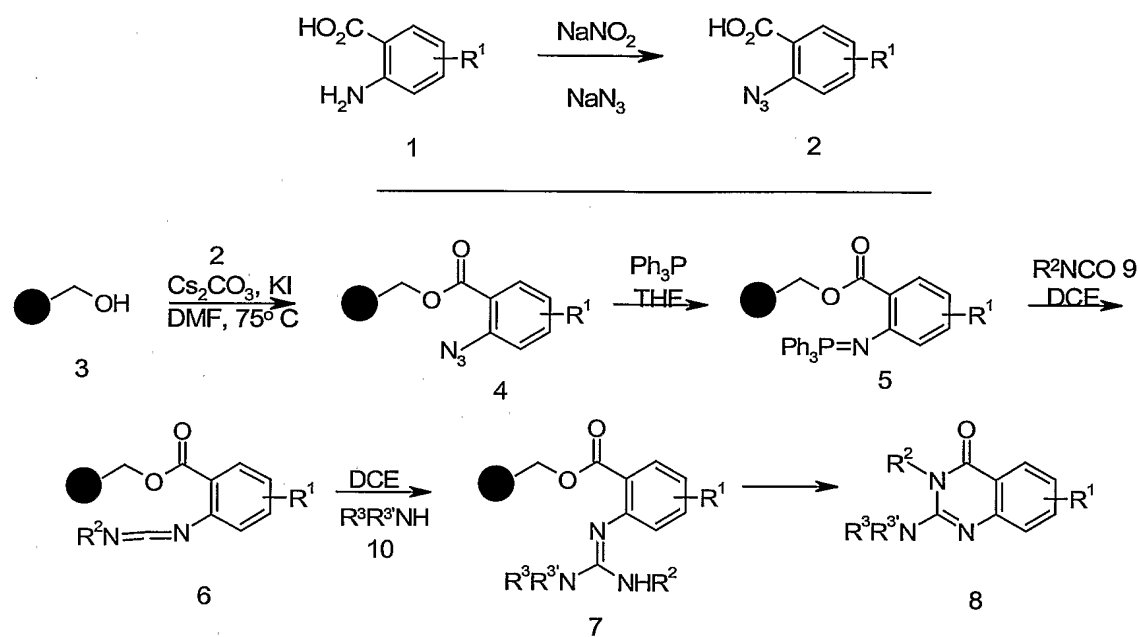
17. A method according to any of claims 13 to 16 wherein the expression of one or more of the genes out of the group comprising ABCA1, ABCG1, ABCG5, ABCG8, apolipoprotein-CI, -CII, -CIV, -E, LPL (lipoprotein lipase), CETP (cholesteryl ester transfer protein) or other genes that positively regulate cholesterol homeostasis is increased upon compound administration.
18. A method according to any of claims 13 to 16 wherein the expression of one or more of the genes out of the group comprising 11- $\beta$ -HSD (11- $\beta$  hydroxysteroid dehydrogenase), PEPCK (phosphoenolpyruvat carboxykinase), G-6-P (glucose-6-phosphatase) is reduced upon compound administration.
19. A method according to any of claims 13 to 16 wherein the expression of one or more of the genes out of the group comprising FAS (Fatty Acid Synthase), SREBP-1c (Sterol-response element binding protein), SCD-1 (Stearoyl-CoA Desaturase), Angiopoietin like protein 3 (Angptl3) or other genes which are relevant for controlling serum triglyceride or glucose levels are not or more weakly increased in liver and or other organs compared to administration of a full agonist like TO901317.
20. A method of blocking in a mammal the cholesterol or fatty acid absorption in the intestine of a mammal in need of such blocking comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound according to claims 1 to 9.
21. A method for treating obesity in a mammal comprising administering a therapeutically effective amount of a compound according to any of claims 1 to 9.
22. A method of modulating a gene whose expression is regulated by the NR1H3 and/or NR1H2 receptor in a mammal comprising administering a therapeutically effective amount of a compound according to claims 1 to 9.



23. A method according to any of claims 13 to 19 wherein the expression of ABCA1 and/or ABCG1 and/or ABCG5 and/or ABCG8 are increased.
24. Use of a compound according to any of claims 1 to 9 in a method according to claims 13 to 23 wherein the mammal is a human
25. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for the prevention or treatment of a NR1H3 and/or NR1H2 receptor protein or NR1H3 and/or NR1H2 receptor protein homologue mediated disease or condition in a mammal wherein the prevention or treatment is directly or indirectly accomplished through the binding of the compound to the NR1H3 and/or NR1H2 receptor protein or NR1H3 and/or NR1H2 receptor protein homologue.
26. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for prevention or treatment of a NR1H3 and/or NR1H2 receptor protein mediated disease or condition wherein the mammal is a human.
27. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for regulating the cholesterol transport system in a mammal by activating the NR1H3 and/or NR1H2 receptor.
28. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for regulating levels of cholesterol, triglyceride, and/or bile acid.
29. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for treating in a mammal atherosclerosis, alzheimer disease, gallstone disease, lipid disorders, inflammatory disorder, type II diabetes, obesity or a cardiovascular disorder.
30. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament capable for blocking in a mammal the cholesterol and/or fatty acid absorption in the intestine.

31. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for treating obesity in a mammal.
32. Use of a compound according to any of claims 1 to 9 for the manufacture of a medicament for modulating a gene whose expression is regulated by the NR1H3 and/or NR1H2 receptor.
33. Use of a compound according to any of claims 1 to 9 in a mammal for the selective up-regulation of one or more genes selected from the group comprising ABCA1, ABCG1, ABCG5, ABCG8, apolipoprotein-CI, -CII, -CIV, -E, LPL (lipoprotein lipase), CETP (cholesteryl ester transfer protein) or other genes that positively regulate cholesterol homeostasis are increased and a weaker regulation of one or more of the genes selected from the group comprising FAS and SREBP-1c or other genes that positively regulate lipogenesis, said compound showing a larger difference in regulation of the two groups of genes when compared with the regulatory behaviour of a full agonist like T0901317 on both groups of genes.
34. Use of a compound according any of claims 1 to 9 in a use according to claims 25 to 33 wherein the mammal is human.

Fig.1



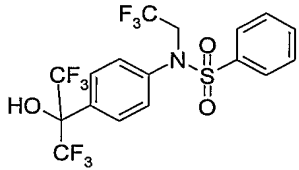
**Fig. 2**

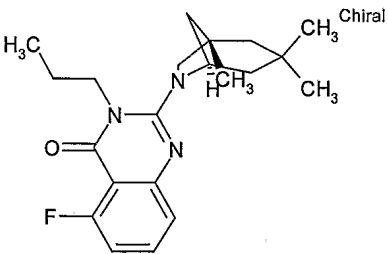
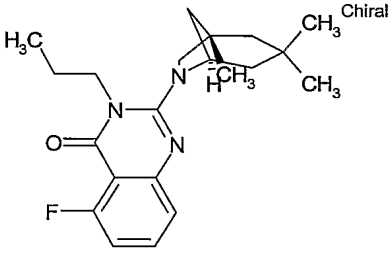
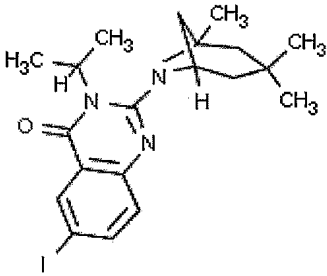
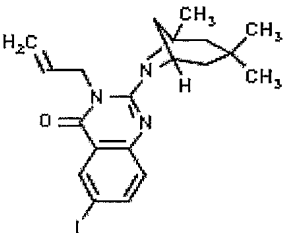
Number of repeats .....	1
plate: GREINER FIA-Plate black 384 well med. binding	
Measurement height .....	3.50 mm
Label technology .....	TR-F Lance
Emission filter name .....	D615
Emission filter slot .....	A1
Emission aperture .....	Normal
Excitation filter .....	D340
Delay .....	50 $\mu$ s
Window time .....	400 $\mu$ s
Cycle .....	1000 $\mu$ s
Light integrator capacitors .....	1
Light integrator ref. level .....	95
Flash energy area .....	High
Flash energy level .....	223
Flash absorbance measurement .....	No
Beam .....	Normal
Label technology .....	TR-F Lance
Emission filter name .....	D665
Emission filter slot .....	A8
Emission aperture .....	Normal
Excitation filter .....	D340
Delay .....	50 $\mu$ s
Window time .....	400 $\mu$ s
Cycle .....	1000 $\mu$ s
Light integrator capacitors .....	1
Light integrator ref. level .....	95
Flash energy area .....	High
Flash energy level .....	223
Flash absorbance measurement .....	No
Beam .....	Normal

**Fig. 3**

<b>Gene definition</b>	<b>NCBI Accession number of gene</b>	<b>Corresponding SEQ ID Number</b>
Liver X receptor beta NR1H2	NM_007121	<b>1</b>
Liver X receptor alpha NR1H3	NM_005693	<b>2</b>
Steroid receptor coactivator 1	U90661	<b>3</b>

**Fig. 4 A**

MOLNAME	MOLECULE STRUCTURE	EC50 AVG [μM]	EFF AV
TO901317	 <p>TO901317</p>	0,1	10

MOLNAME	MOLECULE STRUCTURE	EC50 AVG ( $\mu$ M)	%EF AV
TR1040002211		1,5	50
TR1040002212		4	75
TR1040011382		12	80
TR1040001892		4	50

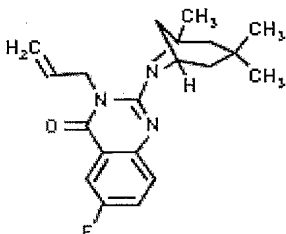
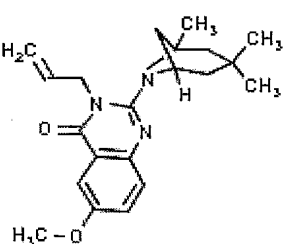
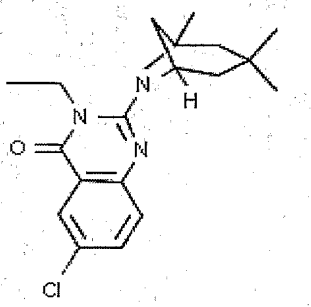
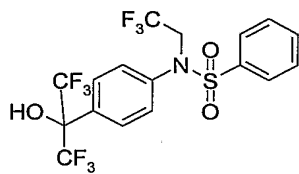
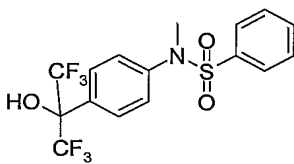
LN0000006662		4	60
LN0000006674		5	20
LN0000007460		5	50



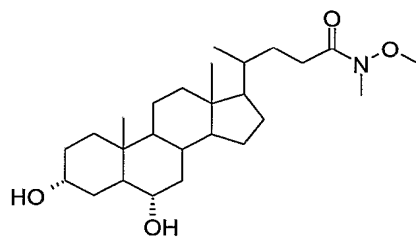
Fig. 5



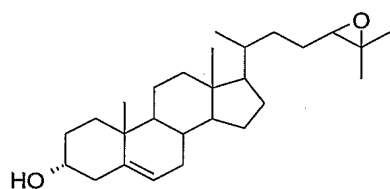
TO901317



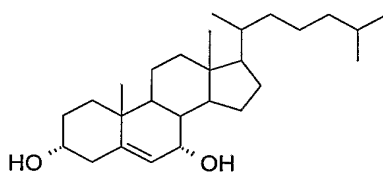
TO314407



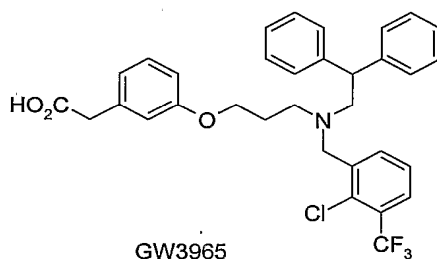
Hypocholamide



24(S),25-Epoxycholesterol



7a-Hydroxycholesterol



GW3965

**Fig. 6**

<b>Protein / Gene Name</b>	<b>NCBI Acccession number of gene</b>	<b>Corresponding SEQ ID number</b>
Liver X receptor beta, LXR $\beta$	NM_007121	<b>1</b>
Liver X receptor alpha, LXR $\alpha$	NM_005693	<b>2</b>
Cholesterol 7 $\alpha$ hydroxylase, Cyp7A1	NM_000780	<b>3</b>
Fatty Acid Synthase FAS	NM_004104	<b>4</b>
Stearyl CoA desaturase, SCD	XM_030447	<b>5</b>
Sterol Response Element Binding Protein 1C, SREBP-1C	NM_004176	<b>6</b>
ATP binding cassette transporter A1; ABCA1	NM_005502	<b>7</b>
ATP binding cassette transporter G1; ABCG1	XM_032950	<b>8</b>
ATP binding cassette transporter G5; ABCG5	NM_022436	<b>9</b>
ATP binding cassette transporter G8; ABCG8	AF324494	<b>10</b>
Apolipoprotein E, apoE	NM_000041	<b>11</b>
Apolipoprotein C-I, apoC-I	NM_001645	<b>12</b>
Apolipoprotein C-II apoC-II	NM_000483	<b>13</b>
Apolipoprotein C-IV, apoC-IV	U32576	<b>14</b>
Lipoprotein Lipase, LPL	M15856	<b>15</b>
Cholesteryl Ester Transfer Protein, CETP	NM_000078	<b>16</b>
Phosphoenolpyruvate carboxykinase 1 (PEPCK)	NM_002591	<b>17</b>
Glucose-6-phosphatase (G6P)	NM_000151	<b>18</b>
Insulin-responsive glucose transporter (GLUT4)	M20747	<b>19</b>
Angiopoietin-like 3, ANGPTL3	NM_01445	<b>20</b>
11-beta Hydroxysteroid dehydrogenase HSD11B1 variant 2	NM_181755	<b>21</b>
11-beta Hydroxysteroid dehydrogenase HSD11B1 variant 1	NM_005525	<b>22</b>

Fig. 7A

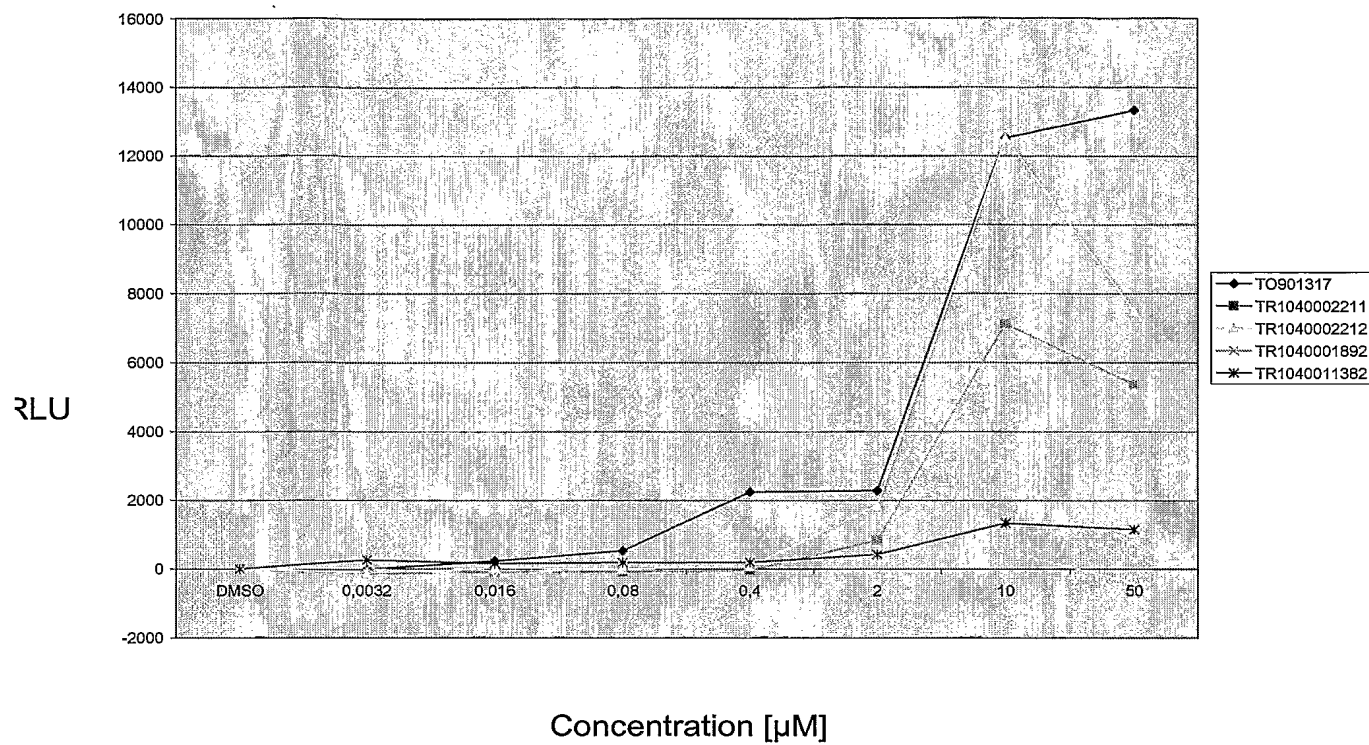


Fig. 7B

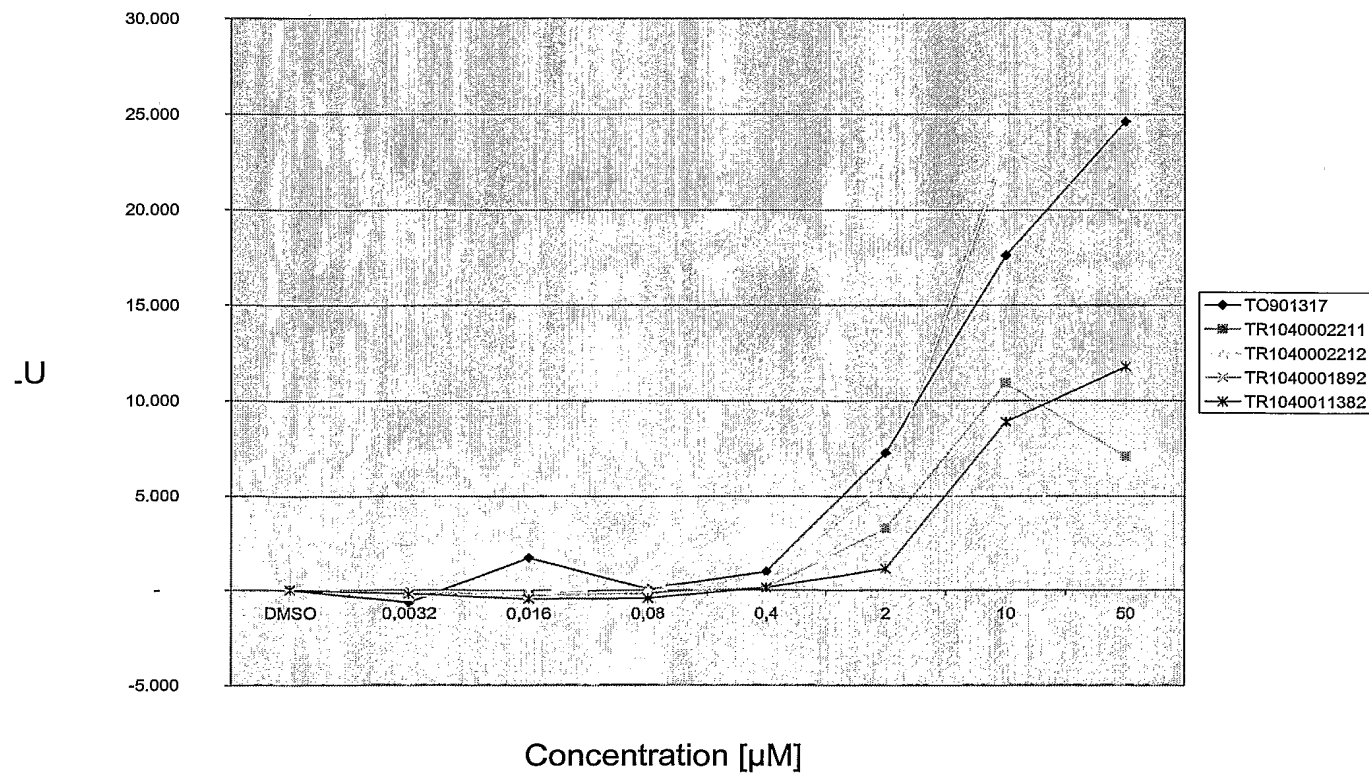


Fig. 8A

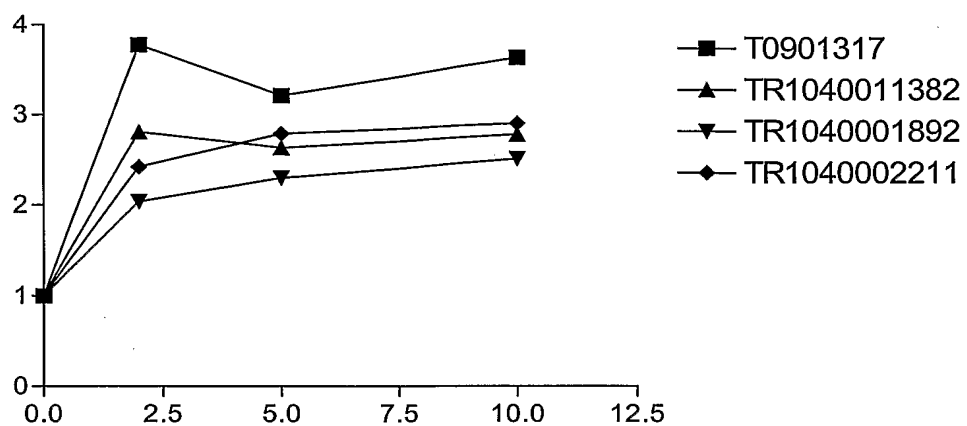


Fig. 8B

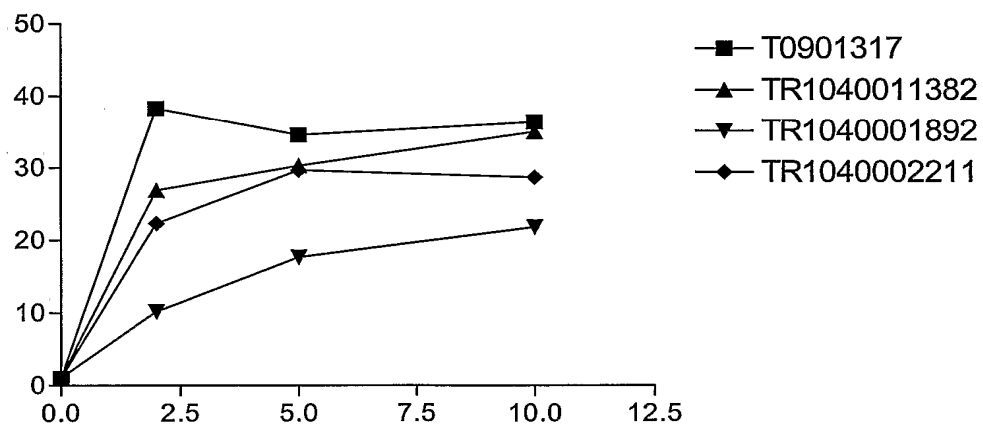


Fig. 8C

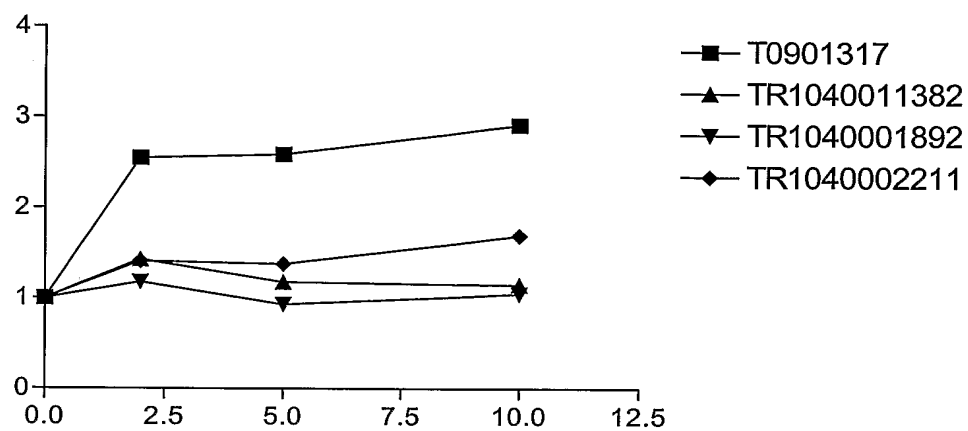
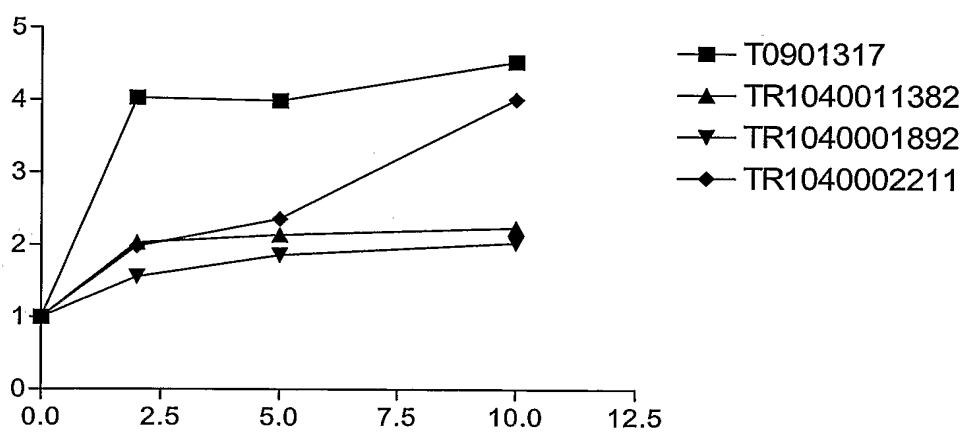
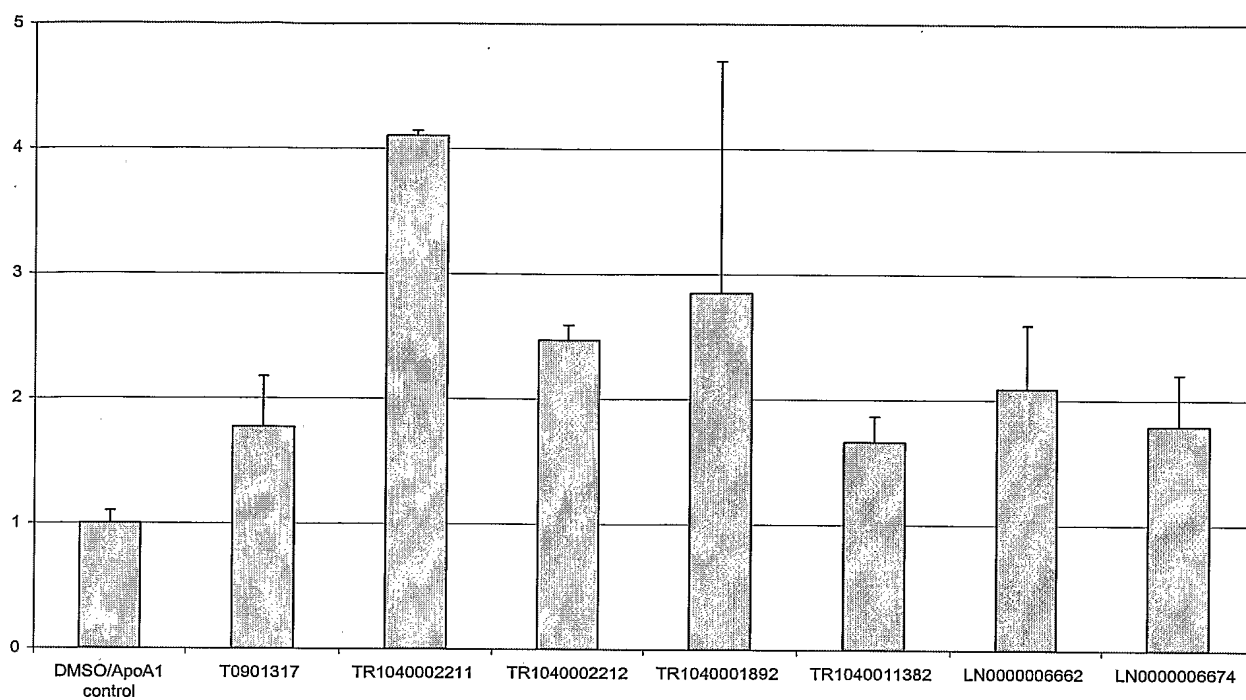
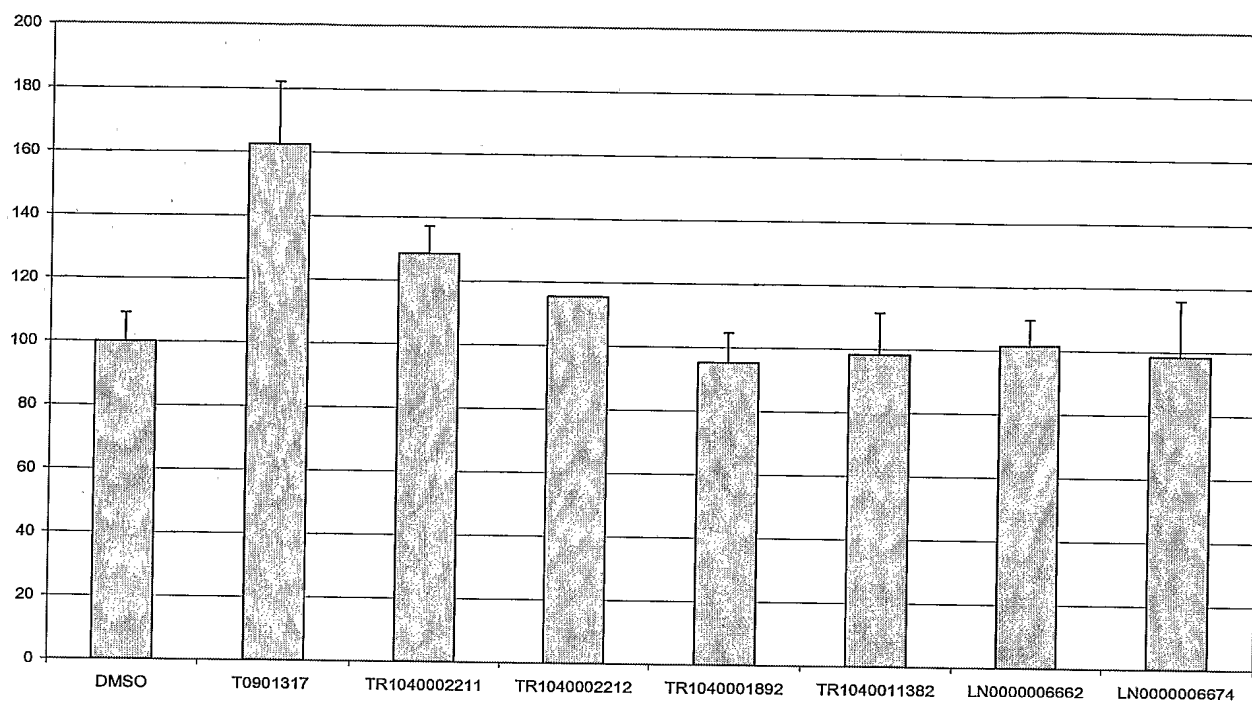


Fig. 8D



**Fig. 9 A**

**Fig. 9 B**

## SEQUENCE LISTING

<110> Phenex Pharmaceuticals AG

<120> Novel selective LXR Nuclear Receptor Binding Compounds with partial agonistic properties

<130> PX62420PC

<160> 23

<170> PatentIn version 3.1

<210> 1

<211> 2010

<212> DNA

<213> Homo sapiens

<400> 1

```

caagaagtgg cgaagttacc tttgagggta tttgagtagc ggcggtgtgt caggggctaa      60
agaggaggac gaagaaaagc agagcaaggg aaccagggc aacaggagta gttcactccg      120
cgagaggccg tccacgagac cccgcgcgc aggcattgagc cccgcccccc acgcattgagc      180
cccgcccccc gctgttgctt ggagaggggc gggacctgga gagaggctgc tccgtgacct      240
caccatgtcc tctcctacca cgagttccct ggatacccc ctgcctggaa atggcccccc      300
tcagcctggc gccccttctt cttcaccac tgtaaaggag gaggggccg agccgtggcc      360
cgggggtccg gaccctgatg tcccaggcac tgataggcc agctcagcct gcagcacaga      420
ctgggtcatc ccagatcccc aagaggaacc agagcgcaag cgaaagaagg gccagcccc      480
gaagatgctg ggccacgagc tttgccgtgt ctgtggggac aaggcctccg gcttcacta      540
caacgtgctc agctgcgaag gctgcaagg cttcttccg cgagtggtg tccgtgggtg      600
ggccaggcgc tatgcctgcc ggggtggcgg aacctgccag atggacgctt tcatgcggcg      660
caagtgccag cagtgccggc tgcgcaagt caaggaggca gggatgagg agcagtgcgt      720
cctttctgaa gaacagatcc ggaagaagaa gattcgaaa cagcagcagc aggagtcaca      780
gtcacagtgc cagtcacctg tggggccgca gggcagcagc agctcagcct ctgggcctgg      840
ggcttcccct ggtggatctg aggcaggcag ccagggtctc ggggaaggcg aggggtgtcca      900
gctaacagcg gctcaagaac taatgatcca gcagttggtg gcggcccaac tgcagtgcaa      960

```



```

caaacgctcc ttctccgacc agcccaaagt cacgccctgg cccctgggcg cagaccccca 1020
gtcccgagat gcccgcagc aacgctttgc ccacttcacg gagctggcca tcatctcagt 1080
ccaggagatc gtggacttcg ctaagcaagt gcctggtttc ctgcagctgg gccgggagga 1140
ccagatcgcc ctctgaagg catccactat cgagatcatg ctgctagaga cagccaggcg 1200
ctacaaccac gagacagagt gtatcacctt cttgaaggac ttcacctaca gcaaggacga 1260
cttcaccgt gcaggcctgc aggtggagtt catcaacccc atcttcgagt tctcggggc 1320
catgcgggcg ctgggcctgg acgacgctga gtacgccctg ctcatcgcca tcaacatctt 1380
ctcggccgac cgGCCAACg tgcaggagcc gggccgcgtg gaggcgttgc agcagcccta 1440
cgtggaggcg ctgctgtcct acacgcgcat caagaggccg caggaccagc tgcgcttccc 1500
gcgcatgctc atgaagctgg tgagcctgcg cacgctgagc tctgtgcact cggagcaggt 1560
cttcgccttg cggctccagg acaagaagct gccgcctctg ctgtcggaga tctgggacgt 1620
ccacgagtga ggggctggcc acccagcccc acagccttgc ctgaccacc tccagcagat 1680
agacgccggc accccttcct ctctctaggg tggaaggggc cctgggagc cctgtagacc 1740
tatcggctct catcccttgg gataagcccc agtccaggtc caggaggctc cctccctgcc 1800
cagcgagtct tccagaagg gtgaaaggg tgcaggctcc gaccactgac ctttccggc 1860
tgccctccct cccagctta cacctcaagc ccagcacgca gcgtaccttg aacagagggga 1920
ggggaggacc catggctctc cccccctagc ccgggagacc aggggccttc ctcttctct 1980
gcttttattt aataaaaata aaaacagaaa 2010

```

<210> 2

<211> 1528

<212> DNA

<213> Homo sapiens

<400> 2

```

cagtgccttg gtaatgacca gggctccaga aagagatgtc cttgtggctg ggggcccctg 60
tgcttgacat tcctcctgac tctgcggtgg agctgtggaa gccaggcgca caggatgcaa 120
gcagccaggc ccagggaggc agcagctgca tcctcagaga ggaagccagg atgccccact 180
ctgctggggg tactgcaggg gtggggctgg aggctgcaga gcccacagcc ctgctacca 240
gggcagagcc ccttcagaa cccacagaga tccgtccaca aaagcggaaa aaggggccag 300
ccccaaaat gctggggaac gagctatgca gcgtgtgtgg ggacaaggcc tcgggcttcc 360
actacaatgt tctgagctgc gagggctgca agggattctt ccgccgcagc gtcataaagg 420
gagcgacta catctgccac agtggcggcc actgccccat ggacacctac atgcgtcgca 480
agtgccagga gtgtcggctt cgcaaatgcc gtcaggctgg catgcgggag gagtgtgtcc 540
tgtcagaaga acagatccgc ctgaagaaac tgaagcggca agaggaggaa caggctcatg 600

```

```

ccacatcctt gccccccagg cgttcctcac ccccccaaat cctgccccag ctcagcccgg 660
aacaactggg catgatcgag aagctcgtcg ctgcccagca acagtgtaac cggcgctcct 720
tttctgaccg gcttcgagtc acgccttggc ccatggcacc agatcccat agccgggagg 780
cccgtcagca gcgctttgcc cacttcactg agctggccat cgtctctgtg caggagatag 840
ttgactttgc taaacagcta cccggcttcc tgcagctcag ccgggaggac cagattgccc 900
tgctgaagac ctctgcgac gaggtgatgc ttctggagac atctcggagg tacaaccctg 960
ggagtgaag tatcaccttc ctcaaggatt tcagttataa ccgggaagac tttgccaaag 1020
cagggtcgca agtgggaattc atcaacccca tcttcgagtt ctccagggcc atgaatgagc 1080
tgcaactcaa tgatgccgag tttgccttgc tcattgctat cagcatcttc tctgcagacc 1140
ggcccaacgt gcaggaccag ctccagggtg agaggctgca gcacacatat gtggaagccc 1200
tgcatgccta cgtctccatc caccatcccc atgaccgact gatgttcca cggatgctaa 1260
tgaaactggg gagcctccg accctgagca gcgtccactc agagcaagtg tttgcactgc 1320
gtctgcagga caaaaagctc ccaccgctgc tctctgagat ctgggatgtg cacgaatgac 1380
tgttctgtcc ccatattttc tgttttcttg gccggatggc tgaggcctgg tggctgcctc 1440
ctagaagtgg aacagactga gaagggcaaa cattcctggg agctgggcaa ggagatcctc 1500
ccgtggcatt aaaagagagt caaagggt 1528

```

&lt;210&gt; 3

&lt;211&gt; 2877

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 3
gtggcatcct tccctttcta atcagagatt ttcttcctca gagatttttg cctagatttg 60
caaaatgatg accacatctt tgattttggg gattgtctata gcagcatgct gttgtctatg 120
gcttatttctt ggaattagga gaaggcaaac ggggtgaacca cctctagaga atggattaat 180
tccatacctg ggctgtgctc tgcaattttg tgccaatcct cttgagttcc tcagagcaaa 240
tcaaaggaaa catggtcatg tttttacctg caaactaatg ggaaaatatg tccatttcat 300
caciaatccc ttgtcatacc ataagggtgt gtgccacgga aaatattttg attggaaaaa 360
atttcacttt gctacttctg cgaaggcatt tgggcacaga agcattgacc cgatggatgg 420
aaataccact gaaaacataa acgacacttt catcaaaacc ctgcagggcc atgccttgaa 480
ttccctcacg gaaagcatga tggaaaacct ccaacgtatc atgagacctc cagtctcctc 540
taactcaaag accgctgcct gggtgacaga agggatgtat tctttctgct accgagtgat 600
gtttgaagct gggatatttaa ctatctttgg cagagatctt acaaggcggg acacacagaa 660
agcacatatt ctaaacaatc ttgacaactt caagcaattc gacaaagtct ttccagccct 720

```

ggtagcaggc	ctccccattc	acatgttcag	gactgcgcac	aatgcccggg	agaaactggc	780
agagagcttg	aggcacgaga	acctccaaaa	gagggaaagc	atctcagaac	tgatcagcct	840
gcgcatgttt	ctcaatgaca	ctttgtccac	ctttgatgat	ctggagaagg	ccaagacaca	900
cctcgtggtc	ctctgggcat	cgcaagcaaa	caccattcca	gcgactttct	ggagtttatt	960
tcaaatgatt	aggaacccag	aagcaatgaa	agcagctact	gaagaagtga	aaagaacatt	1020
agagaatgct	gggtcaaaaag	tcagcttgga	aggcaatcct	atgtgtttga	gtcaagcaga	1080
actgaatgac	ctgccagtat	taaatagtat	aatcaaggaa	tcgctgaggc	tttccagtgc	1140
ctccctcaac	atccggacag	ctaaggagga	tttcactttg	caccttgagg	acggttccta	1200
caacatccga	aaagatagca	tcatagctct	ttaccacacag	ttaatgcact	tagatccaga	1260
aatctacca	gaccctttga	cttttaaata	tgataggtat	cttgatgaaa	acgggaagac	1320
aaagactacc	ttctatttga	atggactcaa	gttaaagtat	tactacatgc	cctttggatc	1380
gggagctaca	atatgtcctg	gaagattggt	cgctatccac	gaaatcaagc	aatttttgat	1440
tctgatgctt	tcttattttg	aattggagct	tatagagggc	caagctaaat	gtccaccttt	1500
ggaccagtcc	cgggcaggct	tgggcatttt	gccgccattg	aatgatattg	aatttaaata	1560
taaattcaag	catttgtgaa	tacatggctg	gaataagagg	acactagatg	atattacagg	1620
actgcagaac	accctcacca	cacagtcctt	ttggacaaat	gcatttagtg	gtggtagaaa	1680
tgattcacca	gggtccaatgt	tgttcaccag	tgcttgcttg	tgaatcttaa	cattttgggtg	1740
acagtttcca	gatgctatca	cagactctgc	tagtgaaaag	aactagtttc	taggagcaca	1800
ataatttggt	ttcatttgta	taagtccatg	aatgttcata	tagccagggg	ttgaagttta	1860
ttattttcaa	aggaaaacac	ctttatttta	ttttttttca	aaatgaagat	acacattaca	1920
gccagggtgtg	gtagcaggca	cctgtagtct	tagctactcg	agaggccaaa	gaaggaggat	1980
ggcttgagcc	caggagtcca	agaccagcct	ggacagctta	gtgagatccc	gtctccgaag	2040
aaaagatatg	tattctaatt	ggcagattgt	tttttcctaa	ggaaactgct	ttatttttat	2100
aaaactgcct	gacaattatg	aaaaaatggt	caaattcacg	ttctagttaa	actgcattat	2160
ttgttgacta	gatgggtggg	ttcttcgggt	gtgatcatat	atcataaagg	atatttcaaa	2220
tgattatgat	tagttatgtc	ttttaataaa	aaggaaatat	ttttcaactt	cttctatatc	2280
caaaattcag	ggcttttaac	atgattatct	tgattttcca	aaaacactaa	aggtgggttt	2340
attttccctt	catgttttaa	cttattgttg	ctgaaaactc	tatgtccggc	tttaactatc	2400
ttctctatat	ttttatttca	ttcacattaa	tgagaagagt	tttctcagag	attaaaaaag	2460
gtagtttttc	tgtcattggt	aaatacacat	tatcactgaa	aaaatgtagc	ttttatgatg	2520
tatgttttaa	agttaaaact	ggatggaaat	agccatttgg	aagctttggg	tatgaaacat	2580
gtggagtgtg	ttaagtgcag	cttgacatta	tgttttatct	aaatgccttt	tatcgctaaa	2640
tgacttgcag	atgaaaaaaa	ctaagggtgac	tcgagtgttt	aaatgcctgt	gtacaacaat	2700
gctttgataa	aatattttta	ggtatgagtt	atcagctcta	tgtcaattga	tattttctgtg	2760

tagtattttat atttaaatta tattttacctt tttgcttatt ttacaaatat taagaaaata 2820  
 ttctaacatt tgataatttt gaaatgattc atctttcaga aataaaaagta tgaatct 2877

<210> 4

<211> 8461

<212> DNA

<213> Homo sapiens

<400> 4

gagggagcca gagagacggc agcggccccc gcctccctct ccgccgcgct tcagcctccc 60  
 gctccgcccgc gctccagcct cgctctccgc cgcccgcacc gccgcccgcg ccctcaccag 120  
 agcagccatg gaggaggtgg tgattgccgg catgtccggg aagctgccag agtcggagaa 180  
 cttgcaggag ttctgggaca acctcatcgg cgggtgtggac atgggtcacgg acgatgaccg 240  
 tcgctggaag gcggggctct acggcctgcc ccggcggtcc ggcaagctga aggacctgtc 300  
 taggtttgat gcctccttct tcggagtcca ccccaagcag gcacacacga tggaccctca 360  
 gctgcggctg ctgctggaag tcacctatga agccatcgtg gacggaggca tcaaccacga 420  
 ttactccga ggaacacaca ctggcgtctg ggtgggcgtg agcggctctg agacctcgga 480  
 ggccctgagc cgagaccccc agacactcgt gggctacagc atgggtgggct gccagcgagc 540  
 gatgatggcc aaccggctct ctttcttctt cgacttcaga gggcccagca tcgcaactgga 600  
 cacagcctgc tcctccagcc tgatggccct gcagaacgcc taccaggcca tccacagcgg 660  
 gcagtgccct gccgccatcg tggggggcat caatgtcctg ctgaagccca acacctccgt 720  
 gcagtctctt aggctgggga tgctcagccc cgagggcacc tgcaaggcct tcgacacagc 780  
 ggggaatggg tactgccgct cggagggtgt ggtggccgtc ctgctgacca agaagtccct 840  
 ggcccggcgg gtgtacgcca ccatcctgaa cgccggcacc aatacagatg gcttcaagga 900  
 gcaaggcgtg accttccccct caggggatat ccaggagcag ctcatccgct cgttgtacca 960  
 gtcggccgga gtggcccctg agtcatttga atacatcgaa gccacaggca caggcaccaa 1020  
 ggtgggcgac cccaggagc tgaatggcat caccgagcc ctgtgcgcca cccgccagga 1080  
 gccgctgctc atcggctcca ccaagtccaa catggggcac ccggagccag cctcggggct 1140  
 ggagccctg gccaaaggct tgctgtccct ggagcacggg ctctgggccc ccaacctgca 1200  
 cttccatagc cccaaccctg agatcccagc gctgttgat gggcggtgc aggtggtgga 1260  
 ccagcccctg ccgctccgtg gcggcaacgt gggcatcaac tcctttggct tcgggggctc 1320  
 caacgtgcac atcatcctga ggcccaacac gcagccgccc ccgcacccg cccacatgc 1380  
 caccctgccc cgtctgctgc gggccagcgg acgcaccct gaggccgtgc agaagctgct 1440  
 ggagcagggc ctccggcaca gccaggacct ggctttcctg agcatgctga acgacatcgc 1500  
 ggctgtcccc gccaccgcca tgcccttccg tggctacgct gtgctgggtg gtgagcgcg 1560

tggcccagag	gtgcagcagg	tggccgctgg	cgagcgcccc	ctctggttca	tctgctctgg	1620
gatgggcaca	cagtggcgcg	ggatggggct	gagcctcatg	cgcttgacc	gcttccgaga	1680
ttccatccta	cgctccgatg	aggctgtgaa	gccattcggc	ctgaagggtgt	cacagctgct	1740
gctgagcaca	gacgagagca	cctttgatga	catcgtccat	tcgtttgtga	gcctgactgc	1800
catccagata	ggcctcatag	acctgctgag	ctgcatgggg	ctgaggccag	atggcatcgt	1860
cggccactcc	ctgggggagg	tggcctgtgg	ctacgccgac	ggctgcctgt	cccaggagga	1920
ggccgtcctc	gctgcctact	ggaggggaca	gtgcatcaaa	gaagcccatc	tcccggcggg	1980
cgccatggca	gccgtgggct	tgtcctggga	ggagtgtaaa	cagcgctgcc	ccccggcggg	2040
ggtgcccgcc	tgccacaact	ccaaggacac	agtcaccatc	tcgggacctc	aggccccggg	2100
gtttgagttc	gtggagcagc	tgaggaagga	gggtgtgttt	gccaaggagg	tgcggaccgg	2160
cggtatggcc	ttccactcct	acttcattga	ggccatcgca	ccccactgc	tgcaggagct	2220
caagaagggtg	atccgggagc	cgaagccacg	ttcagcccgc	tggctcagca	cctctatccc	2280
cgaggcccag	tggcacagca	gcctggcacg	cacgtcctcc	gccgagtaca	atgtcaacaa	2340
cctggtgagc	cctgtgctgt	tccaggaggc	cctgtggcac	gtgcctgagc	acgcggtggt	2400
gctggagatc	gcgccccacg	ccctgctgca	ggctgtcctg	aagcgtggcc	tgaagccgag	2460
ctgcaccatc	atccccctga	tgaagaagga	tcacagggac	aacctggagt	tcttcctggc	2520
cggcatccgg	aggctgcacc	tctcaggcat	cgacgccaac	cccaatgcct	tgttcccacc	2580
tgtggagttc	ccagctcccc	gaggaactcc	cctcatctcc	ccactcatca	agtgggacca	2640
cagcctggcc	tgggacgtgc	cggccgcca	ggacttcccc	aacggttcag	gttccccctc	2700
agccgccatc	tacaacatcg	acaccagctc	cgagtctcct	gaccactacc	tgggtggacca	2760
caccctcgac	ggtcgctcc	tcttccccgc	cactggctac	ctgagcatag	tgtggaagac	2820
gctggccccga	cccctggggc	tgggcgtcga	gcagctgcct	gtgggtgtttg	aggatgtggt	2880
gctgcaccag	gccaccatcc	tgcccaagac	tgggacagtg	tccttgagg	tacggctcct	2940
ggaggcctcc	cgtgccttcg	aggtgtcaga	gaacggcaac	ctggtagtga	gtgggaagggt	3000
gtaccagtgg	gatgaccctg	accccaggct	cttcgaccac	ccggaaagcc	ccacccccaa	3060
ccccacggag	cccctcttcc	tggcccaggc	tgaagtttac	aaggagctgc	gtctgcgtgg	3120
ctacgactac	ggccctcatt	tccagggcct	cctggaggcc	agcctggaag	gtgactcggg	3180
gaggctgctg	tggaaggata	actgggtgag	cttcattggac	accatgctgc	agatgtccat	3240
cctgggctcg	gccaaacacg	gcctgtacct	gccacccgt	gtcaccgcca	tccacatcga	3300
ccctgccacc	cacaggcaga	agctgtacac	actgcaggac	aaggcccaag	tggctgacgt	3360
ggtggtgagc	aggtggctga	gggtcacagt	ggccggaggc	gtccacatct	ccgggctcca	3420
cactgagtcg	gccccgcggc	ggcagcagga	gcagcagggtg	cccatcctgg	agaagttttg	3480
cttcaactccc	cacacggagg	aggggtgcct	gtctgagcgc	gctgccctgc	aggaggagct	3540
gcaactgtgc	aaggggctgg	tgcaggcact	gcagaccaag	gtgaccagc	aggggctgaa	3600

gatggtggtg	cccggactgg	atggggccca	gatcccccg	gacccctcac	agcaggaact	3660
gccccggctg	ttgtcggtg	cctgcaggct	tcagctcaac	gggaacctgc	agctggagct	3720
ggcgcagggtg	ctggcccagg	agaggcccaa	gctgccagag	gaccctctgc	tcagcggcct	3780
cctggactcc	ccggcactca	aggcctgcct	ggacactgcc	gtggagaaca	tgcccagcct	3840
gaagatgaag	gtggtggagg	tgctggccgg	ccacggtcac	ctgtattccc	gcatcccagg	3900
cctgctcagc	ccccatcccc	tgctgcagct	gagctacacg	gccaccgacc	gccaccccca	3960
ggccctggag	gctgcccagg	ccgagctgca	gcagcacgac	gttgcccagg	gccagtggga	4020
tcccgcagac	cctgccccca	gcgccctggg	cagcgccgac	ctcctggtgt	gcaactgtgc	4080
tgtggctgcc	ctcggggacc	cggcctcagc	tctcagcaac	atggtggctg	ccctgagaga	4140
agggggcttt	ctgctcctgc	acacactgct	ccgggggcac	ccctcgggac	atgtggcctt	4200
cctcacctcc	actgagccgc	agtatggcca	gggcatcctg	agccaggacg	cgtgggagag	4260
cctcttctcc	aggggtgtccg	tgcgcttggg	gggcctgaag	aagtccttct	acggctccac	4320
gctcttcctg	tgccgcccgc	ccaccccga	ggacagcccc	atcttcctgc	cggtggacga	4380
taccagcttc	cgctgggtgg	agtctctgaa	gggcatcctg	gctgacgaag	actcttcccg	4440
gcctgtgtgg	ctgaaggcca	tcaactgtgc	cacctcgggc	gtggtgggct	tggtgaactg	4500
tctccgccga	gagcccggcg	gaacgctccg	gtgtgtgctg	ctctccaacc	tcagcagcac	4560
ctcccacgtc	ccggagggtg	acccgggctc	cgcagaactg	cagaagggtgt	tgaggggaga	4620
cctggtgatg	aacgtctacc	gcgacggggc	ctggggggct	ttccgccact	tcctgctgga	4680
ggaggacaag	cctgaggagc	cgacggcaca	tgcccttgtg	agcaccctca	cccgggggga	4740
cctgtcctcc	atccgctggg	tctgctcctc	gctgcgccat	gcccagccca	cctgccctgg	4800
cgcccagctc	tgacaggctc	actacgcctc	cctcaacttc	cgcgacatca	tgctggccac	4860
tggaagctg	tcccctgatg	ccatcccagg	gaagtggacc	tcccaggaca	gcctgctagg	4920
tatggagttc	tcgggcccag	acgccagcgg	caagcgtgtg	atgggactgg	tgcttgccaa	4980
gggcctggcc	acctctgtcc	tgctgtcacc	ggacttcctc	tgggatgtgc	cttccaactg	5040
gacgctggag	gaggcggcct	cggtgcctgt	cgtctacagc	acggcctact	acgcgctggt	5100
ggtgcgtggg	cgggtgcgcc	ccggggagac	gctgctcatc	cactcgggct	cgggcggcgt	5160
gggccaggcc	gccatcgcca	tcgccctcag	tctgggctgc	cgcgctcttc	ccaccgtggg	5220
gtcggctgag	aagcgggctg	acctccaggc	caggttcccc	cagctcgaca	gcaccagctt	5280
cgccaactcc	cgggacacat	ccttcgagca	gcatgtgctg	tggcacacgg	gcgggaaggg	5340
cgttgacctg	gtcttgaact	ccttggcgga	agagaagctg	caggccagcg	tgagggtgctt	5400
ggctacgcac	ggtcgcttcc	tggaatttgg	caaattcgac	ctttctcaga	accacccgct	5460
cggcatggct	atcttcctga	agaacgtgac	attccacggg	gtcctactgg	atgcgttctt	5520
caacgagagc	agtgtgact	ggcgggaggt	gtgggcgctt	gtgcaggccg	gcatccggga	5580
tgggggtggt	cgggccctca	agtgcacggt	gttccatggg	gcccagggtg	aggacgcctt	5640

ccgctacatg	gcccgaagga	agcacattgg	caaagtcgtc	gtgcaggtgc	ttgcggagga	5700
gccggaggca	gtgctgaagg	gggccaaacc	caagctgatg	tcggccatct	ccaagacctt	5760
ctgcccggcc	cacaagagct	acatcatcgc	tggtggtctg	ggtggcttcg	gcctggagtt	5820
ggcgcagtgg	ctgatacagc	gtgggggtgca	gaagctcgtg	ttgacttctc	gctccgggat	5880
ccggacaggc	taccaggcca	agcagggtccg	ccggtggagg	cgccagggcg	tacagggtgca	5940
ggtgtccacc	agcaacatca	gctcactgga	ggggggcccg	ggcctcattg	ccgaggcggc	6000
gcagcttggg	cccgtgggcg	gcgtcttcaa	cctggccgtg	gtcttgagag	atggcttgct	6060
ggagaaccag	accccagagt	tcttccagga	cgtctgcaag	cccaagtaca	gcggcaccct	6120
gaacctggac	aggggtgacc	gagaggcgtg	ccctgagctg	gactactttg	tggtcttctc	6180
ctctgtgagc	tgccggcggtg	gcaatgcggg	acagagcaac	tacggctttg	ccaattccgc	6240
catggagcgt	atctgtgaga	aacgccggca	cgaaggcctc	ccaggcctgg	ccgtgcagtg	6300
gggcgccatc	ggcgacgtgg	gcatttttgt	ggagacgatg	agcaccaacg	acacgatcgt	6360
cagtggcacg	ctgccccagc	gcatggcgctc	ctgcctggag	gtgctggacc	tcttcctgaa	6420
ccagccccac	atggtcctga	gcagctttgt	gctggctgag	aaggctgcgg	cctataggga	6480
cagggacagc	cagcgggacc	tggtggaggc	cgtggcacac	atcctgggca	tccgcgactt	6540
ggctgctgtc	aacctggaca	gctcactggc	ggacctgggc	ctggactcgc	tcatgagcgt	6600
ggaggtgctg	cagacgctgg	agcgtgagct	caacctgggtg	ctgtccgtgc	gcgaggtgctg	6660
gcaactcacg	ctccggaaac	tgcaggagct	gtcctcaaag	gcggatgagg	ccagcgagct	6720
ggcatgcccc	acgcccgaag	aggatggtct	ggcccagcag	cagactcagc	tgaacctgctg	6780
ctccctgctg	gtgaaccg	agggccccac	cctgatgcgg	ctcaactccg	tgagagctc	6840
ggagcggccc	ctgttcctgg	tgcacccaat	cgagggtctc	accaccgtgt	tccacagcct	6900
ggcctcccgg	ctcagcatcc	ccacctatgg	cctgcagtgc	acccgagctg	cgcccccttga	6960
cagcatccac	agcctggctg	cctactacat	cgactgcac	aggcaggtgc	agcccagagg	7020
cccctaccgc	gtggccggct	actcctacgg	ggcctgcgtg	gcctttgaaa	tgtgctccca	7080
gctgcaggcc	cagcagagcc	cagccccac	ccacaacagc	ctcttcctgt	tcgacggctc	7140
gccacctac	gtactggcct	acaccagag	ctaccgggca	aagctgaccc	caggctgtga	7200
ggctgaggct	gagacggagg	ccatatgctt	cttcgtgcag	cagttcacgg	acatggagca	7260
caacaggggtg	ctggaggcgc	tgctgccgct	gaagggccta	gaggagcgtg	tggcagccgc	7320
cgtggacctg	atcatcaaga	gccaccagg	cctggaccgc	caggagctga	gctttgcggc	7380
ccggtccttc	tactacaagc	tgctgcccgc	tgagcagtac	acaccaagg	ccaagtacca	7440
tggcaacgtg	atgctactgc	gcgccaagac	gggtggcgcc	tacggcgagg	acctgggcgc	7500
ggactacaac	ctctcccagg	tatgcgacgg	gaaagtatcc	gtccacgtca	tcgaggggtga	7560
ccaccgcacg	ctgctggagg	gcagcggcct	ggagtccatc	atcagcatca	tccacagctc	7620
cctggctgag	ccacgcgtga	gcgtgcggga	gggctaggcc	cgtgcccccg	cctgccaccg	7680

gaggtcactc caccatcccc accccacccc accccacccc cgccatgcaa cgggattgaa 7740  
 gggtcctgcc ggtgggaccc tgtccggccc agtgccactg ccccccgagg ctgctagacg 7800  
 taggtgttag gcatgtccca cccacccgcc gcctcccacg gcacctcggg gacaccagag 7860  
 ctgccgactt ggagactcct ggtctgtgaa gagccggtgg tgcccgtgcc cgcaggaact 7920  
 gggctgggccc tcgtgcgccc gtgggggtctg cgcttggtct ttctgtgctt ggatttgcat 7980  
 atttattgca ttgctggtag agacccccag gcctgtccac cctgccaaga ctccctcaggc 8040  
 agcgtgtggg tcccgcactc tgccccatt tcccgatgt cccctgcggg cgcgggcagc 8100  
 cacccaagcc tgctggctgc ggccccctct cggccaggca ttggctcagc ccgctgagtg 8160  
 gggggctcgtg ggccagtccc cgaggagctg ggcccctgca caggcacaca gggcccggcc 8220  
 acaccagcg gcccccgca cagccacccg tggggtgctg cccttatgcc cggcgccggg 8280  
 caccaactcc atgtttggtg tttgtctgtg tttgtttttc aagaaatgat tcaaattgct 8340  
 gcttggaattt tgaaatttac tgtaactgtc agtgtacacg tctggacccc gtttcatttt 8400  
 tacaccaatt tggtaaaaat gctgctctca gcctcccaca attaaaccgc atgtgatctc 8460  
 c 8461

<210> 5

<211> 1444

<212> DNA

<213> Homo sapiens

<400> 5

acggtcaccc gttgccagct ctagccttta aattcccggc tcggggacct ccacgcaccg 60  
 cggctagcgc cgacaaccag ctagcgtgca aggcgccgcg gctcagcgcg taccggcggg 120  
 cttcgaaacc gcagtcctcc ggcgaccccg aactccgctc cggagcctca gccccctgga 180  
 aagtgatccc ggcattccgag agccaagatg ccggcccact tgctgcagga cgatatctct 240  
 agctcctata ccaccaccac caccattaca gcgcctccct ccagggtcct gcagaatgga 300  
 ggagataagt tggagacgat gcccctctac ttggaagacg acattcgccc tgatataaaa 360  
 gatgatatat atgacccac ctacaaggat aaggaaggcc caagcccaa ggttgaatat 420  
 gtctggagaa acatcatcct tatgtctctg ctacacttgg gagccctgta tgggatcact 480  
 ttgattccta cctgcaagtt ctacacctgg ctttgggggg tattctacta tttgtcagt 540  
 gccctgggca taacagcagg agctcatcgt ctgtggagcc accgctctta caaagctcgg 600  
 ctgcccctac ggctctttct gatcattgcc aacacaatgg cattccagaa tgatgtctat 660  
 gaatgggctc gtgaccaccg tgcccaccac aagttttcag aaacacatgc tgatcctcat 720  
 aattcccgcac gtggcctttt cttctctcac gtgggttggc tgcttgtgcg caaacaccca 780  
 gctgtcaaag agaaggggag tacgctagac ttgtctgacc tagaagctga gaaactggtg 840



```

atgttccaga ggaggtacta caaacctggc ttgctgatga tgtgcttcat cctgcccacg      900
cttgtgccct ggtatttctg gggtgaaact tttcaaaaaca gtgtgttcgt tgccactttc      960
ttgcgatatg ctgtggtgct taatgccacc tggctggtga acagtgctgc ccacctcttc     1020
ggatatcgtc cttatgacaa gaacattagc ccccgggaga atatcctggg ttcacttgga     1080
gctgtgggtg agggcttcca caactaccac cactcctttc cctatgacta ctctgccagt     1140
gagtaccgct ggcacatcaa cttcaccaca ttcttcattg attgcatggc cgccctcggg     1200
ctggcctatg accggaagaa agtctccaag gccgccatct tggccaggat taaaagaacc     1260
ggagatggaa actacaagag tggctgagtt tgggggtccct caggttcctt tttcaaaaac     1320
cagccaggca gaggttttaa tgtctgttta ttaactactg aataatgcta ccaggatgct     1380
aaagatgatg atgttaaccc attccagtac agtattcttt taaaattcaa aagtattgaa     1440
agcc                                                                    1444

```

&lt;210&gt; 6

&lt;211&gt; 4154

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 6
taacgaggaa cttttcgccg gcgccggggc gcctctgagg ccagggcagg acacgaacgc      60
gcggagcggc ggcggcgact gagagccggg gccgcggcgg cgctccctag gaagggccgt     120
acgagggcggc gggccccggc ggcctcccgg aggaggcggc tgcgccatgg acgagccacc     180
cttcagcgag gcggctttgg agcaggcgct gggcgagccg tgcgatctgg acgcggcgct     240
gctgaccgac atcgaagaca tgcttcagct tatcaacaac caagacagtg acttccctgg     300
cctatttgac ccaccctatg ctgggagtg ggcagggggc acagaccctg ccagccccga     360
taccagctcc ccaggcagct tgtctccacc tcctgccaca ttgagctcct ctcttgaagc     420
cttcctgagc gggccgcagg cagcgccctc acccctgtcc cctccccagc ctgcaccac     480
tccattgaag atgtaccgt ccatgcccgc tttctcccct gggcctggta tcaaggaaga     540
gtcagtgcc a ctgagcatcc tgcagacccc caccacacag cccctgccag gggccctcct     600
gccacagagc ttcccagccc cagccccacc gcagttcagc tccacccttg tgttaggcta     660
ccccagccct ccgggaggct tctctacagg aagccctccc gggaacaccc agcagccgct     720
gcctggcctg cactggctt ccccgccagg ggtcccgccc gtctccttgc acaccaggt     780
ccagagtgtg gtccccagc agctactgac agtcacagct gccccacgg cagccctgt     840
aacgaccact gtgacctgc agatccagca ggtcccggtc ctgctgcagc cccacttcat     900
caaggcagac tcgctgttct tgacagccat gaagacagac ggagccactg tgaaggcggc     960
aggtctcagt cccctggtct ctggcaccac tgtgcagaca gggcctttgc cgaccctggg     1020

```

gagtggcgga accatcttgg caacagtccc actggtcgta gatgcggaga agctgcctat	1080
caaccggctc gcagctggca gcaaggcccc ggcctctgcc cagagccgtg gagagaagcg	1140
cacagcccac aacgccattg agaagcgcta ccgctcctcc atcaatgaca aaatcattga	1200
gctcaaggat ctggtggtgg gcactgaggc aaagctgaat aaatctgctg tcttgcgcaa	1260
ggccatcgac tacattcgct ttctgcaaca cagcaaccag aaactcaagc aggagaacct	1320
aagtctgcgc actgctgtcc acaaaagcaa atctctgaag gatctggtgt cggcctgtgg	1380
cagtggaggg aacacagacg tgctcatgga gggcgtgaag actgaggtgg aggacacact	1440
gacccacccc ccctcggatg ctggctcacc tttccagagc agccccttgt cccttggcag	1500
caggggcagt ggcagcgggtg gcagtggcag tgactcggag cctgacagcc cagtctttga	1560
ggacagcaag gcaaagccag agcagcggcc gtctctgcac agccggggca tgctggaccg	1620
ctcccgccctg gccctgtgca cgctcgtctt cctctgcctg tcttgcaacc ccttggcctc	1680
cttgctgggg gcccgggggc tttccagccc ctgagatacc accagcgtct accatagccc	1740
tgggcgcaac gtgctgggca ccgagagcag agatggccct ggctggggcc agtggctgct	1800
gcccccagtg gtctggctgc tcaatgggct gttggtgctc gtctccttgg tgcttctctt	1860
tgtctacggt gagccagtca cacggcccca ctgagggccc gccgtgtact tctggaggca	1920
tcgcaagcag gctgacctgg acctggcccc gggagacttt gcccaggctg cccagcagct	1980
gtggctggcc ctgcgggcac tgggcccggc cctgcccacc tcccacctgg acctggcttg	2040
tagcctcctc tggaacctca tccgtcacct gctgcagcgt ctctgggtgg gccgctggct	2100
ggcaggcccg gcagggggcc tgcagcagga ctgtgctctg cgagtggatg ctagcgccag	2160
cgcccgagac gcagccctgg tctaccataa gctgcaccag ctgcacacca tggggaagca	2220
cacaggcggg cacctactg ccaccaacct ggcgctgagt gccctgaacc tggcagagtg	2280
tgcaggggat gccgtgtctg tggcgacgct ggccgagatc tatgtggcgg ctgcattgag	2340
agtgaagacc agtctccac gggccttgca ttttctgaca cgcttcttcc tgagcagtgc	2400
ccgccaggcc tgccctggcac agagtggctc agtgccctcct gccatgcagt ggctctgcca	2460
ccccgtgggc caccgtttct tcgtggatgg ggactggctc gtgctcagta ccccatggga	2520
gagcctgtac agcttgggcg ggaaccagct ggacccctg gcccagggtga ctgagctatt	2580
ccgggaacat ctcttagagc gagcactgaa ctgtgtgacc cagcccaacc ccagccctgg	2640
gtcagctgat ggggacaagg aattctcgga tgccctcggg tacctgcagc tgctgaacag	2700
ctgttctgat gctgcggggg ctctgccta cagcttctcc atcagttcca gcatggccac	2760
caccaccggc gtagaccggg tggccaagtg gtgggcctct ctgacagctg tggatgacca	2820
ctggctgcgg cgggatgagg aggcggctga gcggtgtgct ccgctggtgg agcacctgcc	2880
ccgggtgctg caggagtctg agagaccct gcccagggca gctctgcact cttcaaggc	2940
tgcccgggcc ctgctgggct gtgccaaggc agagtctggt ccagccagcc tgaccatctg	3000
tgagaaggcc agtgggtacc tgcaggacag cctggctacc acaccagcca gcagctccat	3060

tgacaaggcc	gtgcagctgt	tcctgtgtga	cctgcttctt	gtggtgcgca	ccagcctgtg	3120
gcggcagcag	cagcccccg	ccccggcccc	agcagcccag	ggcaccagca	gcaggcccca	3180
ggcttccg	cttgagctgc	gtggcttcca	acgggacctg	agcagcctga	ggcggctggc	3240
acagagcttc	cggcccgc	tgcgagggt	gttcctacat	gaggccacgg	cccggctgat	3300
ggcggggg	agccccacac	ggacacacca	gctcctcgac	cgagctctga	ggcggcgggc	3360
agggcccggt	ggcaaaggag	gcgcgggtggc	ggagctggag	ccgcggccca	cgcggcggga	3420
gcacgcggag	gccttgctgc	tggcctcctg	ctacctgccc	cccggcttcc	tgtcggcgcc	3480
cgggcagcgc	gtgggcatgc	tggctgaggc	ggcgcgcaca	ctcgagaagc	ttggcgatcg	3540
ccggctgctg	cacgactgtc	agcagatgct	catgcgcctg	ggcggtgagg	ccactgtcac	3600
ttccagctag	accccggtgc	cccggcctca	gcacccctgt	ctctagccac	tttgggtccc	3660
tgcagcttct	gtcctgcgtc	gaagctttga	aggccgaagg	cagtgaaga	gactctggcc	3720
tccacagttc	gacctgcggc	tgctgtgtgc	cttcgcgggtg	gaaggcccga	ggggcgcgat	3780
cttgacccta	agaccggcgg	ccatgatggt	gctgacctct	ggtggccgat	cggggcactg	3840
caggggcccga	gccattttgg	ggggcccccc	tccttgctct	gcaggcacct	tagtggcttt	3900
tttcctcctg	tgtacaggga	agagaggggt	acatttccct	gtgctgacgg	aagccaactt	3960
ggctttccc	gactgcaagc	agggtctctgc	cccagaggcc	tctctctccg	tcgtgggaga	4020
gagacgtgta	catagtgtag	gtcagcgtgc	ttagcctcct	gacctgaggc	tcctgtgcta	4080
ctttgccttt	tgaaaacttt	attttcatag	attgagaagt	tttgtacaga	gaattaaaaa	4140
tgaaattatt	tata					4154

&lt;210&gt; 7

&lt;211&gt; 10412

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 7

gtaattgcga	gcgagagtga	gtggggccgg	gacccgcaga	gccgagccga	cccttctctc	60
ccgggctgcg	gcagggcagg	gcggggagct	ccgcgcacca	acagagccgg	ttctcagggc	120
gctttgctcc	ttgttttttc	cccggttctg	ttttctcccc	ttctccggaa	ggcttgtaaa	180
ggggtaggag	aaagagacgc	aaacacaaaa	gtggaaaaca	gttaatgacc	agccacggcg	240
tccctgctgt	gagctctggc	cgctgccttc	cagggtctcc	gagccacacg	ctgggggtgc	300
tggctgaggg	aacatggctt	gttggcctca	gctgagggtg	ctgctgtgga	agaacctcac	360
tttcagaaga	agacaaacat	gtcagctgct	gctggaagtg	gcctggcctc	tatttatctt	420
cctgatcctg	atctctgttc	ggctgagcta	cccaccctat	gaacaacatg	aatgccattt	480
tccaaataaa	gccatgccct	ctgcaggaac	acttccttgg	gttcagggga	ttatctgtaa	540

tgccaacaac ccctgtttcc gttacccgac tcctggggag gctcccggag ttgttgga	600
ctttaacaaa tccattgtgg ctgcctgtt ctcatatgct cggaggcttc tttatacag	660
ccagaaagac accagcatga aggacatgcg caaagtcttg agaacattac agcagatcaa	720
gaaatccagc tcaaacttga agcttcaaga tttcctggtg gacaatgaaa ctttctctgg	780
gttcctgtat cacaacctct ctctcccaaa gtctactgtg gacaagatgc tgagggtga	840
tgtcattctc cacaaggat ttttgcaagg ctaccagtta catttgacaa gtctgtgcaa	900
tggatcaaaa tcagaagaga tgattcaact tgggtgacaa gaagtttctg agctttgtgg	960
cctaccaagg gagaaactgg ctgcagcaga gcgagtactt cgttccaaca tggacatcct	1020
gaagccaatc ctgagaacac taaactctac atctcccttc ccgagcaagg agctggctga	1080
agccacaaaa acattgctgc atagtcttgg gactctggcc caggagctgt tcagcatgag	1140
aagctggagt gacatgagc aggaggtgat gtttctgacc aatgtgaaca gctccagctc	1200
ctccaccaa atctaccagg ctgtgtctcg tattgtctgc gggcatcccg agggaggggg	1260
gctgaagatc aagtctctca actggtatga ggacaacaac taaaagccc tctttggagg	1320
caatggcact gaggaagatg ctgaaacctt ctatgacaac tctacaactc cttactgcaa	1380
tgatttgatg aagaatttgg agtctagtcc tctttccgcg attatctgga aagctctgaa	1440
gccgctgctc gttgggaaga tcctgtatac acctgacact ccagccacaa ggcaggcat	1500
ggctgaggtg aacaagacct tccaggaact ggctgtgttc catgatctgg aaggcatgtg	1560
ggaggaactc agcccaaga tctggacctt catggagaac agccaagaaa tggaccttgt	1620
ccggatgctg ttggacagca gggacaatga ccacttttgg gaacagcagt tggatggctt	1680
agattggaca gcccaagaca tcgtggcgtt tttggccaag caccagagg atgtccagtc	1740
cagtaatggt tctgtgtaca cctggagaga agctttcaac gagactaacc aggcaatccg	1800
gaccatatct cgcttcatgg agtgtgtcaa cctgaacaag ctagaaccca tagcaacaga	1860
agtctggctc atcaacaagt ccatggagct gctggatgag aggaagttct gggctggtat	1920
tgtgttcact ggaattactc caggcagcat tgagctgcc catcatgtca agtacaagat	1980
ccgaatggac attgacaatg tggagaggac aaataaaatc aaggatgggt actgggaccc	2040
tggctctcga gctgaccctt ttgaggacat gcggtacgtc tgggggggct tcgcctactt	2100
gcaggatgtg gtggagcagg caatcatcag ggtgctgacg ggcaccgaga agaaaactgg	2160
tgtctatatg caacagatgc cctatccctg ttacgttgat gacatctttc tgcgggtgat	2220
gagccggtca atgcccctct tcatgacgct ggcctggatt tactcagtgg ctgtgatcat	2280
caagggcatc gtgtatgaga aggaggcacg gctgaaagag accatgcgga tcatgggcct	2340
ggacaacagc atcctctggt ttagctggtt cattagtagc ctatttcctc ttcttgtgag	2400
cgctggcctg ctagtgttca tcctgaagtt aggaaacctg ctgccctaca gtgatcccag	2460
cggtggtgtt gtcttctgt cctgttttgc tgtggtgaca atcctgcagt gcttctgat	2520
tagcacactc ttctccagag ccaacctggc agcagcctgt gggggcatca tctacttcac	2580

gctgtacctg	ccctacgtcc	tgtgtgtggc	atggcaggac	tacgtgggct	tcacactcaa	2640
gatcttcgct	agcctgctgt	ctcctgtggc	ttttgggttt	ggctgtgagt	actttgccct	2700
ttttgaggag	cagggcattg	gagtgcagtg	ggacaacctg	tttgagagtc	ctgtggagga	2760
agatggcttc	aatctcacca	cttcggtctc	catgatgctg	tttgacacct	tcctctatgg	2820
ggatgatgacc	tggtacattg	aggctgtctt	tccaggccag	tacggaattc	ccaggccctg	2880
gtattttcct	tgcaccaagt	cctactgggt	tggcgaggaa	agtgatgaga	agagccaccc	2940
tggttccaac	cagaagagaa	tatcagaaat	ctgcatggag	gaggaacca	cccacttgaa	3000
gctgggcgtg	tccattcaga	acctggtaaa	agtctaccga	gatgggatga	aggtggctgt	3060
cgatggcctg	gcactgaatt	tttatgaggg	ccagatcacc	tccttcctgg	gccacaatgg	3120
agcggggaag	acgaccacca	tgtcaatcct	gaccgggttg	ttccccccga	cctcgggcac	3180
cgcctacatc	ctgggaaaag	acattcgctc	tgagatgagc	accatccggc	agaacctggg	3240
ggtctgtccc	cagcataacg	tgctgtttga	catgctgact	gtcgaagaac	acatctgggt	3300
ctatgcccgc	ttgaaagggc	tctctgagaa	gcacgtgaag	gcgagatgg	agcagatggc	3360
cctggatggt	ggtttgccat	caagcaagct	gaaaagcaaa	acaagccagc	tgtcaggtgg	3420
aatgcagaga	aagctatctg	tggccttggc	ctttgtcggg	ggatctaagg	ttgtcattct	3480
ggatgaacct	acagctgggt	tggaccctta	ctcccgcagg	ggaatatggg	agctgctgct	3540
gaaataccga	caaggccgca	ccattattct	ctctacacac	cacatggatg	aagcggacgt	3600
cctggggggac	aggattgcca	tcattctcca	tgggaagctg	tgctgtgtgg	gctcctccct	3660
gtttctgaag	aaccagctgg	gaacaggcta	ctacctgacc	ttggtcaaga	aagatgtgga	3720
atcctccctc	agttcctgca	gaaacagtag	tagcactgtg	tcataacctga	aaaaggagga	3780
cagtgtttct	cagagcagtt	ctgatgctgg	cctgggcagc	gaccatgaga	gtgacacgct	3840
gaccatcgat	gtctctgcta	tctccaacct	catcaggaag	catgtgtctg	aagccccggt	3900
ggatggaagac	atagggcatg	agctgacctt	tgtgctgcca	tatgaagctg	ctaaggaggg	3960
agcctttgtg	gaactctttc	atgagattga	tgaccggctc	tcagacctgg	gcatttctag	4020
ttatggcatc	tcagagacga	ccctggaaga	aatattcctc	aaggtggccg	aagagagtgg	4080
ggatggatgct	gagacctcag	atggtacctt	gccagcaaga	cgaaacaggc	gggccttcgg	4140
ggacaagcag	agctgtcttc	gcccgttcac	tgaagatgat	gctgctgatc	caaatgattc	4200
tgacatagac	ccagaatcca	gagagacaga	cttgctcagt	gggatggatg	gcaaagggtc	4260
ctaccagggtg	aaaggctgga	aacttacaca	gcaacagttt	gtggcccttt	tgtggaagag	4320
actgctaatt	gccagacgga	gtcggaagg	atTTTTTgct	cagattgtct	tgccagctgt	4380
gtttgtctgc	attgcccttg	tgttcagcct	gatcgtgcca	ccctttggca	agtaccccg	4440
cctggaactt	cagccctgga	tgtacaacga	acagtacaca	tttgtcagca	atgatgctcc	4500
tgaggacacg	ggaaccctgg	aactcttaaa	cgccctcacc	aaagaccctg	gcttcggggac	4560
ccgctgtatg	gaaggaaacc	caatcccaga	cacgccctgc	caggcagggg	aggaagagtg	4620

gaccactgcc	ccagttcccc	agaccatcat	ggacctcttc	cagaatggga	actggacaat	4680
gcagaaccct	tcacctgcat	gccagtgtag	cagcgacaaa	atcaagaaga	tgctgcctgt	4740
gtgtccccc	ggggcagggg	ggctgcctcc	tccacaaaga	aaacaaaaca	ctgcagatat	4800
ccttcaggac	ctgacaggaa	gaaacatttc	ggattatctg	gtgaagacgt	atgtgcagat	4860
catagccaaa	agcttaaaga	acaagatctg	ggtgaatgag	tttaggtatg	gcggcttttc	4920
cctgggtgtc	agtaatactc	aagcacttcc	tccgagtcaa	gaagttaatg	atgccatcaa	4980
acaaatgaag	aaacacctaa	agctggccaa	ggacagttct	gcagatcgat	ttctcaacag	5040
cttgggaaga	tttatgacag	gactggacac	caaaaataat	gtcaagggtgt	ggttcaataa	5100
caagggctgg	catgcaatca	gctctttcct	gaatgtcatc	aacaatgcca	ttctccgggc	5160
caacctgcaa	aaggagagaga	accctagcca	ttatggaatt	actgctttca	atcatcccct	5220
gaatctcacc	aagcagcagc	tctcagaggt	ggctctgatg	accacatcag	tggatgtcct	5280
tgtgtccatc	tgtgtcatct	ttgcaatgtc	cttcgtccca	gccagctttg	tcgtattcct	5340
gatccaggag	cgggtcagca	aagcaaaaca	cctgcagttc	atcagtggag	tgaagcctgt	5400
catctactgg	ctctctaatt	ttgtctggga	tatgtgcaat	tacgttgtcc	ctgccacact	5460
ggtcattatc	atcttcatct	gcttccagca	gaagtcctat	gtgtcctcca	ccaatctgcc	5520
tgtgctagcc	cttctacttt	tgctgtatgg	gtgggtcaatc	acacctctca	tgtaccacgc	5580
ctcctttgtg	ttcaagatcc	ccagcacagc	ctatgtgggtg	ctcaccagcg	tgaacctctt	5640
cattggcatt	aatggcagcg	tggccacctt	tgtgctggag	ctgttcaccg	acaataagct	5700
gaataatatc	aatgatatcc	tgaagtccgt	gttcttgatc	ttcccacatt	tttgccctggg	5760
acgagggctc	atcgacatgg	tgaaaaacca	ggcaatggct	gatgccctgg	aaagggtttgg	5820
ggagaatcgc	tttgtgtcac	cattatcttg	ggacttggtg	ggacgaaacc	tcttcgccat	5880
ggccgtggaa	ggggtggtgt	tcttcctcat	tactgttctg	atccagtaca	gattcttcat	5940
caggcccaga	cctgtaaatg	caaagctatc	tcctctgaat	gatgaagatg	aagatgtgag	6000
gcgggaaaga	cagagaattc	ttgatgggtg	aggccagaat	gacatcttag	aaatcaagga	6060
gttgacgaag	atatatagaa	ggaagcggaa	gcctgctggt	gacaggatgt	gcgtgggcat	6120
tcctcctggt	gagtgccttg	ggctcctggg	agttaatggg	gctggaaaat	catcaacttt	6180
caagatgtta	acaggagata	ccactgttac	cagaggagat	gctttcctta	acaaaaatag	6240
tatcttatca	aacatccatg	aagtacatca	gaacatgggc	tactgccctc	agtttgatgc	6300
catcacagag	ctgttgactg	ggagagaaca	cgtggagttc	tttgcccttt	tgagaggagt	6360
cccagagaaa	gaagttggca	aggttggtga	gtgggcgatt	cggaaactgg	gcctcgtgaa	6420
gtatggagaa	aaatatgctg	gtaactatag	tggaggcaac	aaacgcaagc	tctctacagc	6480
catggctttg	atcggcgggc	ctcctgtggt	gtttctggat	gaaccaccca	caggcatgga	6540
tcccaaagcc	cggcggttct	tgtggaattg	tgccctaagt	gttgtcaagg	aggggagatc	6600
agtagtgctt	acatctcata	gtatggaaga	atgtgaagct	ctttgcacta	ggatggcaat	6660

catggtcaat ggaaggttca ggtgccttgg cagtgtccag catctaaaaa ataggtttgg	6720
agatggttat acaatagttg tacgaatagc agggccaac ccggacctga agcctgtcca	6780
ggatttcttt ggacttgcatt ttcctggaag tgttctaaaa gagaaacacc ggaacatgct	6840
acaataccag ctcccatctt cattatcttc tctggccagg atattcagca tcctctccca	6900
gagcaaaaag cgactccaca tagaagacta ctctgtttct cagacaacac ttgaccaagt	6960
atttgtgaac ttgccaagg accaaagtga tgatgaccac ttaaaagacc tctcattaca	7020
caaaaaccag acagtagtgg acgttgcagt tctcacatct tttctacagg atgagaaagt	7080
gaaagaaagc tatgtatgaa gaatcctgtt catacggggg ggctgaaagt aaagaggaaac	7140
tagactttcc ttgccaatgt gtgaagtgtt gtggagaaaa gagccagaag ttgatgtggg	7200
aagaagtaaa ctggatactg tactgatact attcaatgca atgcaattca atgcaatgaa	7260
aacaaaattc cattacaggg gcagtgcctt tgtagcctat gtcttgtatg gctctcaagt	7320
gaaagacttg aatttagttt tttaacctata cctatgtgaa actctattat ggaacccaat	7380
ggacatatgg gtttgaactc acactttttt tttttttttt gttcctgtgt attctcattg	7440
gggttgcaac aataattcat caagtaatca tggccagcga ttattgatca aaatcaaaag	7500
gtaatgcaca tcctcattca ctaagccatg ccatgcccag gagactgggt tcccgggtgac	7560
acatccattg ctggcaatga gtgtgccaga gttattagt ccaagttttt cagaaagttt	7620
gaagcaccat ggtgtgtcat gctcactttt gtgaaagctg ctctgctcag agtctatcaa	7680
cattgaatat cagttgacag aatgggtgcca tgctgtggcta acatcctgct ttgattcccc	7740
ctgataagct gttctgggtg cagtaacatg caacaaaaat gtgggtgtct ccaggcacgg	7800
gaaacttggg tcattgttta tattgtccta tgcttcgagc catgggtcta cagggtcatc	7860
cttatgagac tcttaaatat acttagatcc tggtaagagg caaagaatca acagccaaac	7920
tgctggggct gcaagctgct gaagccaggg catgggatta aagagattgt gcgttcaaac	7980
ctaggggaagc ctgtgcccatt ttgtcctgac tgtctgctaa catggtacac tgcattctaa	8040
gatgtttatc tgacacaagt gtattatttc tggctttttt aattaatcta gaaaatgaaa	8100
agatggagtt gtattttgac aaaaatgttt gtacttttta atgttatttg gaattttaag	8160
ttctatcagt gacttctgaa tccttagaat ggcctctttg tagaacctg tggatatagag	8220
gagtatggcc actgccccac tatttttatt ttcttatgta agtttgcata tcagtcatga	8280
ctagtgccta gaaagcaatg tgatggctcag gatctcatga cattatattt gagtttcttt	8340
cagatcattt aggatactct taatctcact tcatcaatca aatatttttt gagtgtatgc	8400
tgtagctgaa agagtatgta cgtacgtata agactagaga gatattaagt ctcagtacac	8460
ttcctgtgcc atgttattca gctcactggg ttacaaatat aggttgtctt gtggtttag	8520
gagcccactg taacaatact gggcagcctt tttttttttt tttttaattg caacaatgca	8580
aaagccaaga aagtataagg gtcacaagtc taaacaatga attcttcaac agggaaaaca	8640
gctagcttga aaacttgctg aaaaacacaa ctgtgtttta tggcatttag taccttcaaa	8700

```

taattggcctt tgcagatatt ggatacccca tttaatctga cagtctcaaa tttttcatct 8760
cttcaatcac tagtcaagaa aaatataaaa acaacaaata cttccatatg gagcattttt 8820
cagagttttc taaccagtc ttatTTTTtct agtcagtaaa catttgtaaa aatactgttt 8880
cactaatact tactgttaac tgtcttgaga gaaaagaaaa atatgagaga actattgttt 8940
ggggaagttc aagtgatctt tcaatatcat tactaacttc ttccactttt tccagaattt 9000
gaatattaac gctaaaggtg taagacttca gatttcaaat taatctttct atatTTTTt 9060
aatttacaga atattatata acccactgct gaaaaagaaa aaaatgattg ttttagaagt 9120
taaagtcaat attgatttta aatataagta atgaaggcat atttccaata actagtgata 9180
tggcatcggt gcattttaca gtatcttcaa aaatacagaa tttatagaat aatttctcct 9240
catttaatat ttttcaaaat caaagttatg gtttctctcat tttactaaaa tcgtattcta 9300
attcttcatt atagtaaadc tatgagcaac tccttacttc ggttcctctg atttcaaggc 9360
catatTTTaa aaaatcaaaa ggcactgtga actatTTTga agaaaacaca acattTTaat 9420
acagattgaa aggacctctt ctgaagctag aaacaatcta tagttataca tcttcattaa 9480
tactgtgtta cTTTTaaaa tagtaatttt ttacattttc ctgtgtaaac ctaattgtgg 9540
tagaaatttt taccaactct atactcaatc aagcaaaatt tctgtatatt ccctgtggaa 9600
tgtacctatg tgagtttcag aaattctcaa aatacgtgtt caaaaatttc tgctTTTgca 9660
tctttgggac acctcagaaa acttattaac aactgtgaat atgagaaaata cagaagaaaa 9720
taataagccc tctatacata aatgcccagc acaattcatt gttaaaaaaac aaccaaacct 9780
cacactactg tatttcatta tctgtactga aagcaaatgc tttgtgacta ttaaatgttg 9840
cacatcattc attcactgta tagtaatcat tgactaaagc catttgtctg tgTTTTcttc 9900
ttgtggttgt atatatcagg taaaatattt tccaaagagc catgtgtcat gtaatactga 9960
accactttga tattgagaca ttaatttgta cccttggtat tatctactag taataatgta 10020
atactgtaga aatattgctc taattctttt caaaattgtt gcatccccct tagaatgttt 10080
ctatttccat aaggatttag gtatgctatt atcccttctt atacctaag atgaagctgt 10140
ttttgtgctc tttgttcac attggccctc attccaagca ctttacgctg tctgtaatgg 10200
gatctatttt tgcactggaa tatctgagaa ttgcaaaact agacaaaagt ttcacaacag 10260
atTTtctaagt taaatcattt tcattaaaag gaaaaagaa aaaaaatttt gtatgtcaat 10320
aactttatat gaagtattaa aatgcatatt tctatgttgt aatataatga gtcacaaaat 10380
aaagctgtga cagttctgtt ggtctacaga aa 10412

```

&lt;210&gt; 8

&lt;211&gt; 3473

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



<400> 8  
 ttctttccaa gggctctctg gtgaggcccc tgaccttccc aagcctctcc ctgtcttgtg 60  
 aaacctgggc gtgatatacc tcccttttag ggctgctgcg atcatttagg cagattaaac 120  
 ctcataagtg gtttcccata caagaaagat gctagcagtg caacagacag aacacttaac 180  
 tgccctgccct cccgccagga ggtggtcttc caacttttgc ccggagtcta cagagggtgg 240  
 gccctctctg ctggggctcc gggacatggt caggagaggt tggctgtct gtaccgccat 300  
 tctcttggcc agactgtggt gtctggtccc tactcacacc ttcctgtcag agtatccaga 360  
 ggccgcagag tatccacacc ctggctgggt gtactggcta cagatggctg tggctccagg 420  
 tcacctgcgt gcctgggtga tgagaaataa tgtcacaaca aatatcccat ctgcattctc 480  
 tgggacactg acccatgaag agaaagcagt tctcacagtt ttacaggca cagccacagc 540  
 cgtgcatgta cagggtggcag ctttagcttc tgctaaactg gagagctcag tgtttgtgac 600  
 agactgcgtg tcctgcaaaa tcgaaaatgt ctgtgattca gctcttcagg gaaaaaggg 660  
 gccgatgtct ggcctacagg gctcaagcat tgtcatcatg ccccatcca accgtccact 720  
 cgccagtgcg gcatcctgca cgtggtcagt ccaagtccag ggaggggccc atcacctggg 780  
 ggtggtcgct atcagtggca aagtcttgtc agcagctcat ggggcaggaa gggcctatgg 840  
 ttgggggttt cctggcgatc ccatggagga aggatacaag accctcctga aaggaatttc 900  
 cggaagtgc aatagtgggt agttggtggc cattatgggt ccttccgggg ccgggaagtc 960  
 cacgctgatg aacatcctgg ctggatacag ggagacgggc atgaaggggg ccgtcctcat 1020  
 caacggcctg ccccgggacc tgcgtgctt ccggaagggt tcctgctaca tcatgcagga 1080  
 tgacatgctg ctgccgcac tcaactgtgca ggaggccatg atggtgtcgg cacatctgaa 1140  
 gcttcaggag aaggatgaag gcagaagga aatggtcaag gagatactga cagcgtggg 1200  
 cttgctgtct tgcccaaca cgcggaccgg gagcctgtca ggtggtcagc gcaagcgct 1260  
 ggccatcgcg ctggagctgg tgaacaaccc tccagtcatg ttcttcgatg agcccaccag 1320  
 cggcctggac agcgctcct gcttcagggt ggtctcgctg atgaaagggc tcgctcaagg 1380  
 gggctcgctc atcatttgca ccatccacca gccagcgcc aaactcttcg agctgttcga 1440  
 ccagctttac gtcctgagtc aaggacaatg tgtgtaccgg ggaaaagtct gcaatcttgt 1500  
 gccatatattg agggatttg gtctgaactg cccaacctac cacaaccag cagattttgt 1560  
 catggagggt gcatccggcg agtacggtga tcagaacagt cggctggtga gagcggttcg 1620  
 ggagggcagc tgtgactcag accacaagag agacctcggg ggtgatgccg aggtgaaccc 1680  
 ttttcttttg caccggccct ctgaagaggt aaagcagaca aaacgattaa aggggttgag 1740  
 aaaggactcc tcgtccatgg aaggctgcca cagcttctct gccagctgcc tcacgcagtt 1800  
 ctgcatcctc ttcaagagga ccttcctcag catcatgagg gactcggtcc tgacacacct 1860  
 gcgcatcacc tcgcacattg ggatcggcct cctcattggc ctgctgtact tggggatcgg 1920  
 gaacgaagcc aagaaggctc tgagcaactc cggcttcctc ttcttctcca tgctgttcct 1980

catgttcgcg gccctcatgc ctactgttct gacattttccc ctggagatgg gagtcttttct	2040
tcgggaacac ctgaactact ggtacagcct gaaggcctac tacctggcca agaccatggc	2100
agacgtgccc tttcagatca tggtcccagt ggcctactgc agcatcgtgt actggatgac	2160
gtcgcagccg tccgacgccg tgcgctttgt gctgtttgcc gcgctgggca ccatgacctc	2220
cctgggtggca cagtccctgg gcctgctgat cggagccgcc tccacgtccc tgcaggtggc	2280
cactttcgtg ggcccagtga cagccatccc ggtgctcctg ttctcgggggt tcttcgtcag	2340
cttcgacacc atccccacgt acctacagtg gatgtcctac atctcctatg tcaggtatgg	2400
gttcgaaggg gtcacacctc ccatctatgg cttagaccgg gaagatctgc actgtgacat	2460
cgacgagacg tgccacttcc agaagtcgga ggccatcctg cgggagctgg acgtggaaaa	2520
tgccaagctg tacctggact tcatcgtact cgggattttc ttcatctccc tccgcctcat	2580
tgccatatttt gtcctcaggt acaaaatccg ggcagagagg taaaacacct gaatgccagg	2640
aaacaggaag attagacact gtggccgagg gcacgtctag aatcgaggag gcaagcctgt	2700
gccccgaccga cgacacagag actcttctga tccaaccctc agaaccgcgt tgggtttgtg	2760
ggtgtctcgt gctcagccac tctgcccagc tgggttgat cttctctcca ttccccttc	2820
tagctttaac taggaagatg taggcagatt ggtggttttt ttttttttaa catacagaat	2880
tttaaatacc acaactgggg cagaatttaa agctgcaaca cagctggtga tgagaggctt	2940
cctcagtcca gtcgctcctt agcaccaggc accgtgggtc ctggatgggg aactgcaagc	3000
agcctctcag ctgatggctg cacagtcaga tgtctggtgg cagagagtcc gagcatggag	3060
cgattccatt ttatgactgt tgtttttcac attttcatct ttctaagggtg tgtctctttt	3120
ccaatgagaa gtcatttttg caagccaaaa gtcgatcaat cgcattcatt ttaagaaatt	3180
ataccttttt agtacttgct gaagaatgat tcagggtaaa tcacatactt tgtttagaga	3240
ggcgaggggt ttaaccgagt caccagctg gtctcataca tagacagcac ttgtgaagga	3300
ttgaatgcag gttccagggt gagggaagac gtggacacca tctccactga gccatgcaga	3360
cattttttaa agctatacaa aaaattgtga gaagacattg gccaaactct tcaaagtctt	3420
tctttttcca cgtgcttctt attttaagcg aaatatattg tttgtttctt cct	3473

&lt;210&gt; 9

&lt;211&gt; 2740

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

aagtcaccagt cctgctgtcc caagggactc cggggtcagg tggagcaggc agggcagtct	60
gccacgggct ccccaactga agccactctg gggaggggtcc ggccaccaga aaatttgccc	120
agctttgctg cctgttggcc atgggtgacc tctcatcttt gacccccgga ggggtccatgg	180

gtctccaagt aaacagagggc tcccagagct ccctggaggg ggctcctgcc accgccccgg	240
agcctcacag cctgggcatc ctccatgcct cctacagcgt cagccaccgc gtgaggccct	300
ggtagggacat cacatcttgc cggcagcagt ggaccaggca gatcctcaaa gatgtctcct	360
tgtacgtgga gagcgggcag atcatgtgca tcctaggaag ctcaggctcc gggaaaacca	420
cgctgctgga cgccatgtcc gggaggctgg ggcgcgcggg gaccttcctg ggggagggtgt	480
atgtgaacgg ccgggcgctg cgccgggagc agttccagga ctgcttctcc tacgtcctgc	540
agagcgacac cctgctgagc agcctcaccg tgcgcgagac gctgcactac accgcgctgc	600
tggccatccg ccgcggcaat cccggctcct tccagaagaa ggtggaggcc gtcattggcag	660
agctgagtct gagccatgtg gcagaccgac tgattggcaa ctacagcttg gggggcattt	720
ccacgggtga gcggcgccgg gtctccatcg cagcccagct gctccaggat cctaagggtca	780
tgctgtttga tgagccaacc acaggcctgg actgcatgac tgctaatacag attgtcgtcc	840
tcctggtgga actggctcgc aggaaccgaa ttgtggttct caccattcac cagccccgtt	900
ctgagctttt tcagctcttt gacaaaattg ccattcctgag cttcgagag ctgattttct	960
gtggcacgcc agcggaaatg cttgattttct tcaatgactg cggttaccct tgtcctgaac	1020
attcaaacc ttttgacttc tatatggacc tgacgtcagt ggatacccaa agcaaggaac	1080
gggaaataga aacctccaag agagtccaga tgatagaatc tgcctacaag aaatcagcaa	1140
tttgtcataa aactttgaag aatattgaaa gaatgaaaca cctgaaaacg ttaccaatgg	1200
ttcctttcaa aaccaaagat tctcctggag ttttctctaa actgggtgtt ctcctgagga	1260
gagtgacaag aaacttggtg agaaataagc tggcagtgat tacgcgtctc cttcagaatc	1320
tgatcatggg tttgttcctc cttttcttcg ttctgcgggt ccgaagcaat gtgctaaagg	1380
gtgctatcca ggaccgcgta ggtctccttt accagtttgt gggcgccacc ccgtacacag	1440
gcatgctgaa cgctgtgaat ctgtttcccg tgctgcgagc tgtcagcgac caggagagtc	1500
aggacggcct ctaccagaag tggcagatga tgctggccta tgcactgcac gtcctcccct	1560
tcagcgttgt tgccaccatg attttcagca gtgtgtgcta ctggacgctg ggcttacatc	1620
ctgaggttgc ccgatttgga tatttttctg ctgctctctt ggccccccac ttaattggtg	1680
aattttctaac tcttgctgcta cttggtatcg tccaaaatcc aaatatagtc aacagtgtag	1740
tggctctgct gtccattgctg ggggtgcttg ttggatctgg attcctcaga aacatacaag	1800
aaatgcccac tcctttttaa atcatcagtt attttacatt ccaaaaatat tgcagtgaga	1860
ttcttgtagt caatgagttc tacggactga atttcacttg tggcagctca aatgtttctg	1920
tgacaactaa tccaatgtgt gccttcactc aaggaattca attcattgag aaaacctgcc	1980
caggtgcaac atctagattc acaatgaact ttctgatttt gtattcattt attccagctc	2040
ttgtcatcct aggaatagtt gttttcaaaa taagggatca tctcattagc aggtagttaa	2100
agccatggct gggaaaatgg aagtgaagct gccgactgtg catgactgct ctgaacgtct	2160
gaaatgagag tgccatgtat ttctttcttg acaggacatc tcaagtcttt taaccattaa	2220

gactccat	ttt	gtgcctc	ttg	gatccaagca	ggccttgaat	gcaatggaag	tggtttatag	2280
tcccttgctc	ttacaacttg	cagggacatg	tggttat	tttg	gaaattgtga	ctgagcggac	2340	
ccaagaatgt	aaataatatt	cataaaccta	tgggagactc	gtgtgactat	tttttttcct	2400		
tgttctaggc	acagaaaaaa	ataggtcagc	ttaaaaatat	gtttacattg	gataaaggat	2460		
taggcaaaaa	taaaatgttt	caaggattcc	tgaccataag	tgacagagaa	agagagttgt	2520		
gggttttagat	gaagcaaggt	tatcatgcag	aattgggtaa	gaatgcttct	gttcctggaa	2580		
gaccagagt	taaatgcaga	tgtccacacg	aggggtcgga	gttacctgat	cacatcgaga	2640		
gagtgcctggg	cagatggatg	gtgagcacca	ctgctacaga	gcacccagtg	attttactga	2700		
ggattaaaaat	aaaaaacctg	aggaatgggc	tcaacagtga	2740				

&lt;210&gt; 10

&lt;211&gt; 2679

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

ctccaggaaa	cagagtgaag	acactggccc	tggcaggcag	cagctgggtc	taagagagct	60
gcagcccagg	gtcacagacc	tgtggggccc	atggccggga	aggcggcaga	ggagagaggg	120
ctgccgaaag	gggccactcc	ccaggatacc	tcgggcctcc	aggatagatt	gttctcctct	180
gaaagtgaca	acagcctgta	cttcacctac	agtggccagc	ccaacaccct	ggaggtcaga	240
gacctcaact	gccagggtgga	cctggcctct	caggtccctt	ggtttgagca	gctgggtcag	300
ttcaagatgc	cctggacatc	tcccagctgc	cagaattctt	gtgagctggg	catccagaac	360
ctaagcttca	aagtgagaag	tgggcagatg	ctggccatca	tagggagctc	aggttgtggg	420
agagcctcct	tgctagatgt	gatcactggc	cgagggtcacg	gcggcaagat	caagtcaggc	480
cagatctgga	tcaatgggca	gccagctcg	cctcagctgg	tgaggaagtg	tgtggccac	540
gtgcgccagc	acaaccagct	gtccccaac	ttgactgtgc	gagagacctt	ggccttcatt	600
gccagatgc	ggctgcccag	aaccttctcc	caggcccagc	gtgacaaaag	ggtggaggac	660
gtgatcgcg	agctgcggct	taggcagtgc	gctgacaccc	gcgtgggcaa	catgtacgtg	720
cgggggttgt	cggggggtga	gcgcaggaga	gtcagcattg	gggtgcagct	cctgtggaac	780
ccaggaatcc	ttattctcga	cgaaccacc	tctgggctcg	acagcttcac	agccacaac	840
ctggtgaaga	ccttgtccag	gctggccaaa	ggcaaccggc	tggtgctcat	ctccctccac	900
cagcctcgct	ctgacatctt	caggctgttt	gatctgggtc	tcctgatgac	gtctggcacc	960
cccatctact	taggggaggc	ccagcacatg	gtccagtatt	tcacagccat	cggctacccc	1020
tgtcctcgct	acagcaatcc	tgctgacttc	tatgtggacc	tgaccagcat	tgacaggcgc	1080
agcagagagc	aggaattggc	caccagggag	aaggctcagt	cactcgcagc	cctgtttcta	1140

gaaaaagtgc gtgacttaga tgacttttcta tggaaagcag agacgaagga tcttgacgag 1200  
 gacacctgtg tggaaagcag cgtgacccca ctagacacca actgcctccc gagtcctacg 1260  
 aagatgcctg gggcggtgca gcagtttacg acgctgatcc gtcgtcagat ttccaacgac 1320  
 ttccgagacc tgccccacct cctcatccat ggggcggagg cctgtctgat gtcaatgacc 1380  
 atcggccttc tctatTTTTg ccatgggagc atccagctct ccttcatgga tacagccgcc 1440  
 ctcttgttca tgatcgggtg tctcatccct ttcaacgtca ttctggatgt catctccaaa 1500  
 tgttactcag agagggcaat gctttactat gaactggaag acgggctgta caccactggt 1560  
 ccatatttct ttgccaagat cctcggggag cttccggagc actgtgccta catcatcatc 1620  
 tacgggatgc ccacctactg gctggccaac ctgaggccag gcctccagcc cttcctgctg 1680  
 cacttcctgc tgggtgtggct ggtggtcttc tgttgacgga ttatggccct ggccgccgcg 1740  
 gccctgctcc ccaccttcca catggcctcc ttcttcagca atgccctcta caactccttc 1800  
 tacctcgccg ggggcttcat gataaacttg agcagcctgt ggacagtgcc cgcgtggatt 1860  
 tccaaagtgt cttcctgctg gtggtgtttt gaagggctga tgaagattca gttcagcaga 1920  
 agaacttata aaatgcctct cggaacctc accatcgcg tctcaggaga taaaatcctc 1980  
 agtgccatgg agctggactc gtaccctctc tacgccatct acctcatcgt cattggcctc 2040  
 agcgggtggct tcatggtcct gtactacgtg tccttaaggt tcatcaaaca gaaaccaagt 2100  
 caagactggt gattcacgcc agacgtctgc ccgctggtgg gggacctgag cagacccttc 2160  
 aactgcactc cctcctcagg agcccccttc tggggacagt gaggacaatg accctacaga 2220  
 tgctcagcta catccggccc aggggtgctgc ggtggcacag accagccaca ggatggcagt 2280  
 agaataaaga cagtcgaaag ggatttctgc tctactggcag gagactgcga tgactgggag 2340  
 aaaacctgca ctcggtggca cctacaacgt tgctaattta ttcccttttg atatgcattt 2400  
 atataggcaa ctcgatatag gatgggagca aactaggaat gaattgggta gctagactgt 2460  
 gcaggaattg ttggaacctg gagggaacaa taacagtacc tagcagattt ggcttcatct 2520  
 tccaggggcc ccactcctg tggtagacca ccatcaatac agaaagtgc ctaagatgta 2580  
 ccagcaagat gccatccctt ctttttgtgt ggggtcatgg gctccaaaag ccaacgtgaa 2640  
 caattaaaaa tgtattgagc atctaaaaaa aaaaaaaaaa 2679

<210> 11

<211> 1156

<212> DNA

<213> Homo sapiens

<400> 11

cgcagcggag gtgaaggacg tccttcccca ggagccgact ggccaatcac aggcaggaag 60  
 atgaagggtc tgtgggctgc gttgctggc acattcctgg caggatgcca ggccaagggtg 120

gagcaagcgg	tggagacaga	gccggagccc	gagctgcgcc	agcagaccga	gtggcagagc	180
ggccagcgct	gggaactggc	actgggtcgc	ttttgggatt	acctgcgctg	ggtgcagaca	240
ctgtctgagc	aggtgcagga	ggagctgctc	agctcccagg	tcacccagga	actgagggcg	300
ctgatggacg	agaccatgaa	ggagttgaag	gcctacaaat	cggaaactgga	ggaacaactg	360
accccggtgg	cggaggagac	gcgggcacgg	ctgtccaagg	agctgcaggc	ggcgcaggcc	420
cggctgggcg	cggacatgga	ggacgtgtgc	ggccgccttg	tgcagtaccg	cggcgagggtg	480
caggccatgc	tcggccagag	caccgaggag	ctgcgggtgc	gcctcgcctc	ccacctgcgc	540
aagctgcgta	agcggctcct	ccgcgatgcc	gatgacctgc	agaagcgcct	ggcagtgtac	600
caggccgggg	cccgcgaggg	cgccgagcgc	ggcctcagcg	ccatccgcga	gcgcctgggg	660
cccctggtgg	aacaggggccg	cgtgcggggc	gccactgtgg	gctccctggc	cggccagccg	720
ctacaggagc	gggcccaggc	ctggggcgag	cggctgcgcg	cgcggatgga	ggagatgggc	780
agccggaccc	gcgaccgcct	ggacgagggtg	aaggagcagg	tggcggagggt	gcgcgccaag	840
ctggaggagc	aggcccagca	gatacgccctg	caggccgagg	ccttccaggc	ccgcctcaag	900
agctggttcg	agcccctggt	ggaagacatg	cagcgccagt	gggccgggct	ggtggagaag	960
gtgcaggctg	ccgtgggcac	cagcgccgcc	cctgtgcca	gcgacaatca	ctgaacgccg	1020
aagcctgcag	ccatgcgacc	ccacgccacc	ccgtgcctcc	tgctccgcg	cagcctgcag	1080
cgggagaccc	tgtccccgcc	ccagccgtcc	tcctgggggtg	gaccctagtt	taataaagat	1140
tcaccaagtt	tcacgc					1156

&lt;210&gt; 12

&lt;211&gt; 417

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 12

acctcccaac	caagccctcc	agcaaggatt	caggagtgcc	cctcgggcct	cgccatgagg	60
ctcttcctgt	cgtccccggt	cctggtggtg	gttctgtcga	tcgtcttgga	aggcccagcc	120
ccagcccagg	ggaccccaga	cgtctccagt	gccttgata	agctgaagga	gtttggaaac	180
acactggagg	acaaggctcg	ggaactcatc	agccgcatca	aacagagtga	actttctgcc	240
aagatgcggg	agtggttttc	agagacatth	cagaaagtga	aggagaaact	caagattgac	300
tcatgaggac	ctgaagggtg	acatccagga	ggggcctctg	aaatttccca	caccccagcg	360
cctgtgctga	ggactcccgc	catgtggccc	caggtgccac	caataaaaat	cctaccg	417

&lt;210&gt; 13

&lt;211&gt; 753

&lt;212&gt; DNA

<213> Homo sapiens

<400> 13  
 gttgtggctg tggagcggaa gtgggtctca accactataa atcctctctg tgcccgtccg 60  
 gagctggtga ggacagcctg ccagagtctg gtctctggac actatgggca cagcactcct 120  
 cccagctctg tttcttgtcc tcctggtatt gggatttgag gtccagggga cccaacagcc 180  
 ccagcaagat gagatgccta gcccgaacctt cctcacccag gtgaaggaat ctctctccag 240  
 ttactgggag tcagcaaaga cagccgcccc gaacctgtac gagaagacat acctgcccgc 300  
 tgtagatgag aaactcaggg acttgtagag caaaagcaca gcagccatga gcacttacac 360  
 aggcatTTTT actgaccaag ttctttctgt gctgaaggga gaggagtaac agccagaccc 420  
 cccatcagtg gacaagggga gagtccccta ctcccctgat ccccagggtt cagactgagc 480  
 tcccccttcc cagtagctct tgcatcctcc tcccaactct agcctgaatt cttttcaata 540  
 aaaaatacaa ttcaagttgc ttctcatgga tggcactgct tttctgagga ctcaagggcc 600  
 aagatggagg ggctgactca gtccagccaa catttaatga gcacctactt tatgtatgga 660  
 gctctaacc atgggtccat ggggaataaag cagtgaatag taacaataaa taatcgtaac 720  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 753

<210> 14

<211> 5950

<212> DNA

<213> Homo sapiens

<400> 14  
 gaattccagg ggacatctgt tgcgcaccta ctgtgctcac gctggggcct ctgggatgga 60  
 caggattctg ccaaggcaga catctgggtc aagacagtcc tgcacagttg ttcaggttgt 120  
 ggccaagggtt gcgtttgcag atttgccatg taaaaataca ggatgctcag ttacatttga 180  
 atttcagatt aatagcaaaa aaaactttttt ggtataattc tgaaatattt catgggacat 240  
 atttatacta aaacgtcatg cactgttgat ttgaaattca aatgtaattg ggctcctct 300  
 atttcgtctg gcaagcgtag acaaaagaat ccagtccagg ccaggcgcag tggctcaagc 360  
 ctgtaatccc agcacttttg gaggccgagg cgggcggatc acgaggtcag gagatcgaga 420  
 ccatcctggc taacacggtg aaaccccgtc tctactaaaa atacaaaaaa ttagctgggc 480  
 gtggtggcgg gtgcctgtag tcccagctac tcgggagggt gaggcaggag aatggcgtga 540  
 acctgggagg cggagcttgc agtgagccga gatcgcgcca cagcactcca gcctgggcga 600  
 cagagccaga ctctgtctca aaaaaaaaaa aaagaatcca gtccatagtc ccctgagcca 660  
 tgtgccctgg ggtgcagctg ggtccttcag gagaaaaatg ctctatttct ggcaactggga 720

ccgagcctga	tgtgggTTTT	ttgtttggTTT	ttgtttgtTgt	tattgtTTTT	gagacaaggt	780
ctcgctccac	cacacccggc	taatTTTTgt	atttttagta	gagacggggT	ttcactacgt	840
tggccaggct	ggtcttgaac	tcctgacctc	aagtgagccg	cctgcctcgg	cctcccaaag	900
tgctgggatt	acaggtggga	gccaccgccc	tggccctggg	cctgatgttg	atgacctcct	960
actatgtgca	cctgcagctc	tcctgcatag	gcctcagccg	tcctgcatga	ggacactggg	1020
aggcaggTgc	tccctatcaa	ccccgtgtta	cagttgaaca	aactgagccc	cagaaaagaa	1080
aacgtatttg	cccaggTcac	acggtgaaga	agtgagggat	tcgaagccca	ggTccatctg	1140
aagccagagt	cacccagagg	agaaagagtt	ggaattgaga	actcaaggaa	tgcttggaag	1200
tgatcgggct	cgagcccacc	taggaagaaa	cagaggctgg	agacatgaga	ctgtgttgct	1260
atttcctctc	atcaaccctt	gggccctatt	gaggccctac	cacaagcctg	gccctgcagc	1320
ccagtgacta	ggagaaatta	gacacaagat	aataataaca	gcaatgatct	TTTTTTTTTT	1380
tctgagacgg	agtcttgctc	tttcgcccag	gctggactgc	agtggcgcgga	tctcggtca	1440
atgcaagctc	cacctccag	gttcacgcca	ttctcctgcc	tcagcctccc	gagtagctag	1500
gactacaggc	gcctgccacc	acgcctggct	aatttttcat	atttttagta	gagatggggT	1560
ttcaccgtgt	tagccaagat	ggtctcaatc	tcctgacctc	gtgatccgcc	tgccctcgcc	1620
tcccaaagtg	ttgggggttac	aggcatgagc	caccgcgcct	ggccaacagc	aatgatcttt	1680
gagcacctat	attgccagtc	tccacggtaa	gagctttctt	cattttttgt	tttgttttgt	1740
ttcaagacag	agtcttgctc	tgTcaccag	gctggagtgc	agtggTgtga	tcgcggtca	1800
ctgcagcctt	cacttcccgg	gtTcaagcca	ttctcctgcc	tcagcctccc	aagtagctgg	1860
gattacaggc	acgcatcact	acttctggct	aatttttgta	tttttagtag	ggacagggtt	1920
tttcaccatg	ttggccaggT	tggtctcaaa	ctcctggcct	catatgatct	gccacctcg	1980
gcctcccaaa	gtgctgggat	tacaggcgTg	agccactgcg	cctttctttg	tatttgTtca	2040
agtaatatatac	tgaatatatgt	actgtgcctc	ccacttttatg	gaggaggaaa	ctgaggccag	2100
caaatgaggc	tgtcatggga	ggtggagaca	ggatttgaac	ctgcctcagt	gcaggaggct	2160
caagagcctc	tgtcttctct	cagggcactg	tgTgggaggg	tgagaaggag	ggaggccac	2220
agaggcatga	cctctgattg	ccactgtcac	ctgggccctg	ctctctgaag	tctctgccaa	2280
gcgggggaggt	ggccggggga	gggccctgct	ctgtgcagcc	ttccctcccc	cggcccgcag	2340
agttgagcac	agagggacag	aggcacggaa	ccccagaaa	tgtccctcct	cagaaacagg	2400
ctccaggccc	tgccTgccct	gtgcctctgc	gtgctggTcc	tggcctgcat	tgggggtgag	2460
aagaagtggg	tggagggatg	tggggccac	acctggTggg	tgTgagtgtg	gctgtgtgtc	2520
ctgtggctct	gtagccacgt	gagacatgag	tacggagtgt	gtgcgtttca	tggcgtgcgt	2580
atgcatgtgc	gtgtcgggga	gtgtgtgtgt	cggtggctga	gagtgaagtg	tgaatgtcac	2640
attggTacaa	actgggatca	tctgtgtgtg	tgcacgtgcg	tgcgtggaag	tgggagtatg	2700
cagtcgtggT	aaaaaagtgc	atgtctgtgt	gcatatgtgt	atttgtgtgc	acctgtctct	2760



ctgtggggta	tgtgtgtgca	aaatatattga	gtgtgtggac	atgtgtgagg	gggtgagtgt	2820
gtgctggtgt	gtacgtctgt	gttttgcata	tgcatttttt	tttttttttt	ttagacggag	2880
tctcactctg	tcacccaggc	tggagtgcag	tggtagcagt	ggtgcatct	tggctcactg	2940
catcatccgc	ctacccgttt	caagggattc	tcctgcctca	gtcttcagag	tatttgggac	3000
tacagacaca	cgccaccatg	cctggctaata	tttttttttt	tgagacggag	tctcgctctg	3060
ttacccaggc	tggagtgcag	tggcgtgata	ttggctcact	gcaagctccg	cctcccgggt	3120
tcacgccatt	ctcctgcctc	agcctcccga	gtagctggga	ctacaggagc	ccaccaccac	3180
gcctggctaa	ttttttgtat	ttttactaga	gacgggggtt	cgccgtgtta	gccaggatgg	3240
tctccatata	ctgacctcgt	gatccgcctg	cctcggcctt	ccaaagtgct	aggattatag	3300
gcgtgagcca	ctgcgcctgg	ccaatgcctg	gctaattttt	ttatatattt	ggtagagaca	3360
gggttttgcc	atgttgccca	ggctgggtct	gaaatcctga	cctcagggtga	tccgcccgcc	3420
ttggcctccc	aaagtgtctg	gattacaggc	atgagccacc	acgcccggcc	atgtacttta	3480
tgttaaaaatg	ggatcatatt	ctagatcagc	attatccagt	agaaatttaa	atttttaata	3540
cagggccagg	cacggtggct	catgcctgta	atcccagcac	tttcggaggc	cgaggcgggt	3600
ggatcgcaag	gtcaggagat	ttgagatcat	cctggctaac	agatgggtaa	aaacccatct	3660
ctactaaaaa	tacaaaaaat	tagccatgca	tgggtggcatg	cgctgtagt	cccagctact	3720
cgggaggctg	aggccggaga	atcacttgaa	cccgggaggc	agaggttgca	gtgagccgag	3780
atcgcgccac	tgcattccaa	cctgggtgac	agagcgagac	tccgtctgaa	aaaaaaaaaa	3840
aatttaacac	gtatgtagac	aatgtgcaag	gcaccattcc	atgtgcatcg	tatgtagtaa	3900
ctcttaattc	tcacgataac	cctgaggtag	atattattac	cccgttctac	aaaaggagaa	3960
acagtcctgg	ggagacagga	taagtcaccg	gccaaggcac	acagctagct	acatgtggcc	4020
cccgcgtgac	ggctgggtct	tgtaggcgag	gctttgtcca	gatgcgtggg	tagaagggtct	4080
ggcccggaaa	gaggaactga	cagcaaggct	aagccaatgt	ctgcccctgg	gggcagaaaag	4140
tcacctcctg	ctctccctcc	actgtccaca	gaggtagctc	agacagggtg	ggggtcacag	4200
gagaacgaag	ggagaagggg	gtagttcctg	ggcagcaaaa	tcagggtggg	aagggaggca	4260
tcagaggatg	gcaattagag	aggccattag	aggggaacca	caggcagaca	gggtgacagg	4320
agggactact	gacacaagg	gaagagatgg	cccagccgga	cggggtggct	cacatctgta	4380
atcccagcat	tttgggagcc	cgagggtggg	ggatcacttg	aggtcaggag	ttcgaggccc	4440
caacatggca	aaacccccat	tcttctaaaa	atacaaaaat	tagccgggca	tgatggcaga	4500
tgctgtaat	ccctgctact	cgggaggctg	aggcaggaaa	attgcctgaa	tccaggagggt	4560
ggaggttgca	atgagacgag	atcatgacac	tgcactccac	cctgggcaac	agagcaagag	4620
actgactctg	tctcataaaa	aaaaagaaaa	aagaaaaaaa	aaaagagatg	gctgatgggt	4680
aaagaggggt	tagcgggtcag	gggacacata	agggtaaagg	caggaggcaa	gaggactggc	4740
agggggctgc	ccctgggcca	ccgggagcga	cacaggatga	gcatggaggg	aaaggagaaa	4800

```

ggggattcta ggggtcccagc ctacccaagt tgccctctgg ttccacctag catgccagcc 4860
agaggcccag gaaggaaccc tgagcccccc accaaagcta aagatgagtc gctggagcct 4920
ggtgagggggc aggatgaagg agctgctgga gacagtgggtg aacaggacca gagacgggtg 4980
gcaatggttc tgggtgagggg gtgctggcct ggggtggtggg aggggactcc tgggtctgag 5040
ggaggagggg ctggggcctg gacccctgag tctcaggag gaggaagggtg tgggagtggg 5100
gctgtgaccc ctaggtcttg gaggagtgga ggggttagagc tgagagcagg aactcctagg 5160
tcacagagag gagcggataa atggggcaga gaacacctgg ggagagctgg ggcctccact 5220
gtgatgtcct ctctcctgta ggagcccag caccttccgg ggcttcatgc agacctacta 5280
tgacgaccac ctgagggacc tgggtccgct caccaaggcc tggttcctcg aatccaaaga 5340
cagcctcttg aagaagaccc acagcctgtg cccaggtctt gtctgtgggg acaaggacca 5400
gggttaaaat gttcataaaa gccaggtgtg gttgtggcgg gtgcctgtag tcccagctac 5460
tcaggagggt gaggtaggat gatggcttga gcccaggagt tcgagaccag cctgggcaac 5520
acagcgagat ctcttggggg taaaacaaaa agaaaaaaaa aagttcatac ttctccaata 5580
aataaagtct cacctgtgtc cctgtctgga tccttcccca gtgtggccag aaaaaaacc 5640
acccactgc ctcccaggaa tcaatgagta gaagaggtga cacctgatgg ggaaggaaga 5700
gtaggagggt cggaagggt atcaaggaat aacaccctat tgtgggcttg cggagaatgg 5760
gggacttcaa ggcgtgtcag tttcaggagg gtgagggcag gagcgtgggt ggagtcagca 5820
ggtccccatg atggccctca ctgagagctt cgcccttgtc tcctacaagc tctgactcca 5880
ttcccagtgg gcaccagca cctccaaccc ctccacagcc cccaaccag cctctgtcgg 5940
aggcgaattc 5950

```

&lt;210&gt; 15

&lt;211&gt; 3549

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 15
cccctcttcc tcctcctcaa gggaaagctg ccacttcta gctgccctgc catccccttt 60
aaagggcgac ttgctcagcg ccaaaccgcg gctccagccc tctccagcct ccggctcagc 120
cggctcatca gtcggtccgc gccttgtagc tcctccagag ggacgcgccc cgagatggag 180
agcaaagccc tgctcgtgct. gactctggcc gtgtgggtcc agagtctgac cgctcccgc 240
ggaggggttg ccgccgccga ccaaagaaga gattttatcg acatcgaaag taaatttgcc 300
ctaaggaccc ctgaagacac agctgaggac acttgccacc tcattcccgg agtagcagag 360
tccgtggcta cctgtcattt caatcacagc agcaaacct tcattgggtgat ccatggctgg 420
acggtaacag gaatgtatga gagttgggtg caaaacttg tggccgccct gtacaagaga 480

```

gaaccagact	ccaatgtcat	tgtggtggac	tggctgtcac	gggctcagga	gcattaccca	540
gtgtccgcgg	gctacaccaa	actggtggga	caggatgtgg	cccggtttat	caactggatg	600
gaggaggagt	ttaactaccc	tctggacaat	gtccatctct	tgggatacag	ccttgagacc	660
catgctgctg	gcattgcagg	aagtctgacc	aataagaaag	tcaacagaat	tactggcctc	720
gatccagctg	gacctaactt	tgagtatgca	gaagccccga	gtcgtctttc	tcctgatgat	780
gcagattttg	tagacgtctt	acacacattc	accagagggg	cccctggtcg	aagcattgga	840
atccagaaac	cagttgggca	tgttgacatt	tacccgaatg	gaggtacttt	tcagccagga	900
tgtaacattg	gagaagctat	ccgcgtgatt	gcagagagag	gacttggaga	tgtggaccag	960
ctagtgaagt	gctcccacga	gcgctccatt	catctcttca	tcgactctct	gttgaatgaa	1020
gaaaatccaa	gtaaggccta	caggtgcagt	tccaaggaag	cctttgagaa	agggctctgc	1080
ttgagttgta	gaaagaaccg	ctgcaacaat	ctgggctatg	agatcaataa	agtcagagcc	1140
aaaagaagca	gcaaaatgta	cctgaagact	cgttctcaga	tgccctacaa	agtcttccat	1200
taccaagtaa	agattcattt	ttctgggact	gagagtghaa	cccataccaa	tcaggccttt	1260
gagatttctc	tgtatggcac	cgtggccgag	agtgagaaca	tcccattcac	tctgcctgaa	1320
gtttccacaa	ataagaccta	ctccttccta	atttacacag	aggtagatat	tggagaacta	1380
ctcatgttga	agctcaaatg	gaagagtgat	tcatacttta	gctggtcaga	ctggtggagc	1440
agtcccggct	tcgccattca	gaagatcaga	gtaaaagcag	gagagactca	gaaaaagggtg	1500
atcttctgtt	ctagggagaa	agtgtctcat	ttgcagaaaag	gaaaggcacc	tgcggtattt	1560
gtgaaatgcc	atgacaagtc	tctgaataag	aagtcaggct	gaaactgggc	gaatctacag	1620
aacaaagaac	ggcatgtgaa	ttctgtgaag	aatgaagtgg	aggaagtaac	ttttacaaaa	1680
catacccagt	gtttggggtg	tttcaaaagt	ggattttcct	gaatattaat	cccagcccta	1740
cccttgtag	ttatttttag	agacagtctc	aagcactaaa	aagtggctaa	ttcaatttat	1800
ggggtatagt	ggccaaatag	cacatcctcc	aacgttaaaa	gacagtggat	catgaaaagt	1860
gctgttttgt	cctttgagaa	agaaataatt	gtttgagcgc	agagtataat	aaggctcctt	1920
catgtggcgt	attgggccat	agcctataat	tggttagaac	ctcctatttt	aattggaatt	1980
ctggatcttt	cggactgagg	ccttctcaaa	ctttactcta	agtctccaag	aatacagaaa	2040
atgcttttcc	gcggcacgaa	tcagactcat	ctacacagca	gtatgaatga	tgttttagaa	2100
tgattccctc	ttgctattgg	aatgtggtcc	agacgtcaac	caggaacatg	taacttggag	2160
agggacgaag	aaagggtctg	ataaacacag	aggtttttaa	cagtccctac	cattggcctg	2220
catcatgaca	aagttacaaa	ttcaaggaga	tataaaatct	agatcaatta	attcttaata	2280
ggctttatcg	tttattgctt	aatccctctc	tcccccttct	tttttgtctc	aagattatat	2340
tataataatg	ttctctgggt	agggtgtgaa	aatgagcctg	taatcctcag	ctgacacata	2400
atttgaatgg	tgcagaaaaa	aaaaagatac	cgtaatttta	ttattagatt	ctccaaatga	2460
ttttcatcaa	tttaaaatca	ttcaatatct	gacagttact	cttcagtttt	aggcttacct	2520

tggtcatgct	tcagttgtac	ttccagtgcg	tctcttttgt	tcctggcttt	gacatgaaaa	2580
gataggtttg	agttcaaatt	ttgcattgtg	tgagcttcta	cagatttttag	acaaggaccg	2640
tttttactaa	gtaaaagggg	ggagagggtc	ctgggggtgga	ttcctaagca	gtgcttgtaa	2700
accatcgcgt	gcaatgagcc	agatggagta	ccatgagggt	tgttatttgt	tgtttttaac	2760
aactaatcaa	gagtgagtga	acaactattt	ataaactaga	tctcctattt	ttcagaatgc	2820
tcttctacgt	ataaatatga	aatgataaag	atgtcaaata	tctcagaggc	tatagctggg	2880
aacccgactg	tgaaagtatg	tgatatctga	acacatacta	gaaagctctg	catgtgtgtt	2940
gtccttcagc	ataattcgga	agggaaaaca	gtcgatcaag	ggatgtattg	gaacatgtcg	3000
gagtagaaat	tgttcctgat	gtgccagaac	ttcgaccctt	tctctgagag	agatgatcgt	3060
gcctataaat	agtaggacca	atgttgtgat	taacatcatc	aggcttgga	tgaattctct	3120
ctaaaaataa	aatgatgtat	gatttgttgt	tggcatcccc	tttattaatt	cattaaattt	3180
ctggatttgg	gttgtgacct	aggggtgcatt	aacttaaaag	attcactaaa	gcagcacata	3240
gcactgggaa	ctctggctcc	gaaaaacttt	gttatatata	tcaaggatgt	tctggcttta	3300
cattttattt	attagctgta	aatacatgtg	tggatgtgta	aatggagctt	gtacatattg	3360
gaaaggatcat	tgtggctatc	tgcatttata	aatgtgtggt	gctaactgta	tgtgtcttta	3420
tcagtgatgg	tctcacagag	ccaactcact	cttatgaaat	gggctttaac	aaaacaagaa	3480
agaaacgtac	ttaactgtgt	gaagaaatgg	aatcagcttt	taataaaaatt	gacaacattt	3540
tattaccac						3549

&lt;210&gt; 16

&lt;211&gt; 1790

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 16

gtgaatctct	ggggccagga	agaccctgct	gcccgaaga	gcctcatggt	ccgtgggggc	60
tgggcgga	tacatatag	ggctccaggc	tgaacggctc	gggccactta	cacaccactg	120
cctgataacc	atgctggctg	ccacagtcct	gaccctggcc	ctgctgggca	atgcccatgc	180
ctgctccaaa	ggcacctcgc	acgaggcagg	catcgtgtgc	cgcatcacca	agcctgccct	240
cctgggtgtt	aaccacgaga	ctgccaaggt	gatccagacc	gccttccagc	gagccagcta	300
cccagatatc	acgggcgaga	aggccatgat	gctccttggc	caagtcaagt	atgggttgca	360
caacatccag	atcagccact	tgtccatcgc	cagcagccag	gtggagctgg	tggaagccaa	420
gtccattgat	gtctccattc	agaacgtgtc	tgtggtcttc	aaggggaccc	tgaagtatgg	480
ctacaccact	gcctgggtgg	tgggtattga	tcagtccatt	gacttcgaga	tcgactctgc	540
cattgacctc	cagatcaaca	cacagctgac	ctgtgactct	ggtagagtgc	ggaccgatgc	600

ccctgactgc tacctgtctt tccataagct gctcctgcat ctccaagggg agcgagagcc 660  
 tgggtggatc aagcagctgt tcacaaattt catctccttc accctgaagc tggtcctgaa 720  
 gggacagatc tgcaaagaga tcaacgtcat ctctaacatc atggccgatt ttgtccagac 780  
 aagggctgcc agcatccttt cagatggaga cattgggggtg gacatttccc tgacaggtga 840  
 tcccgtcatc acagcctcct acctggagtc ccatcacaag ggtcatttca tctacaagaa 900  
 tgtctcagag gacctcccc tccccacctt ctgcgccaca ctgctggggg actcccgcat 960  
 gctgtacttc tggttctctg agcgagtctt ccactcgctg gccaaggtag ctttccagga 1020  
 tggccgcctc atgctcagcc tgatgggaga cgagttcaag gcagtgtgtg agacctgggg 1080  
 cttcaacacc aaccaggaaa tcttccaaga ggttgtcggc ggcttcccc gccaggccca 1140  
 agtcaccgtc cactgcctca agatgcccaa gatctcctgc caaaacaagg gagtcgtggt 1200  
 caattcttca gtgatggtga aattcctctt tccacgccca gaccagcaac attctgtagc 1260  
 ttacacattt gaagaggata tcgtgactac cgtccaggcc tcctattcta agaaaaagct 1320  
 cttcttaagc ctcttggtt tccagattac accaaagact gtttccaact tgactgagag 1380  
 cagctccgag tccatccaga gcttcctgca gtcaatgatc accgctgttg gcatccctga 1440  
 ggtcatgtct cggctcgagg tagtgtttac agccctcatg aacagcaaag gcgtgagcct 1500  
 cttcgacatc atcaaccctg agattatcac tcgagatggc ttcctgctgc tgcagatgga 1560  
 ctttggttct cctgagcacc tgctgggtgga tttcctccag agcttgagct agaagtctcc 1620  
 aaggaggtcg ggatggggct tgtagcagaa ggcaagcacc aggctcacag ctggaaccct 1680  
 ggtgtctcct ccagcgtggt ggaagtggg ttaggagtac ggagatggag attggctccc 1740  
 aactcctccc tatectaaag gccactggc attaaagtgc tgtatccaag 1790

<210> 17

<211> 2688

<212> DNA

<213> Homo sapiens

<400> 17

tttaaagctg ggaggttctg ccaccaagca cggccttccc actgggaaca caaacttgct 60  
 ggcgggaaga gcccggaaag aaacctgtgg atctcccttc gagatcatcc aaagagaaga 120  
 aaggtagcct cacattcgtg ccccttagca gcactctgca gaaatgcctc ctgagctgca 180  
 aaacggcctg aacctctcgg ccaaagttgt ccaggggaagc ctggacagcc tgccccaggc 240  
 agtgagggag tttctcgaga ataacgtga gctgtgtcag cctgatcaca tccacatctg 300  
 tgacggctct gaggaggaga atgggcggct tctgggccag atggaggaag agggcatcct 360  
 caggcggctg aagaagtatg acaactgctg gttggctctc actgacccca gggatgtggc 420  
 caggatcgaa agcaagacgg ttatcgtcac ccaagagcaa agagacacag tgcccatccc 480

caaaacaggc	ctcagccagc	tcggctcgctg	gatgtcagag	gaggattttg	agaaagcggt	540
caatgccagg	ttcccagggt	gcatgaaagg	tcgcaccatg	tacgtcatcc	cattcagcat	600
ggggccgctg	ggctcgccctc	tgtcaaagat	cggcatcgag	ctgacggatt	caccctacgt	660
ggtggccagc	atgcggatca	tgacgcggat	gggcacgccc	gtcctggaag	cactgggcga	720
tggggagttt	gtcaaagtcc	tccatttctgt	ggggtgccct	ctgcctttac	aaaagccttt	780
ggtcaacaac	tggccctgca	acccggagct	gacgctcatc	gcccacctgc	ctgaccgcag	840
agagatcatc	tcctttggca	gtgggtacgg	cgggaactcg	ctgctcggga	agaagtgctt	900
tgctctcagg	atggccagcc	ggctggccaa	ggaggaaggg	tggctggcag	agcacatgct	960
ggttctgggt	ataaccaacc	ctgaggggtga	gaagaagtac	ctggcggccg	cattttccag	1020
cgctgcggg	aagaccaacc	tggccatgat	gaacccagc	ctccccgggt	ggaaggttga	1080
gtgcgtcggg	gatgacattg	cctggatgaa	gtttgacgca	caaggtcatt	taagggccat	1140
caaccagaa	aatggctttt	tcgggtgtcg	tcctgggact	tcagtgaaga	ccaaccccaa	1200
tgccatcaag	accatccaga	agaacacaat	ctttaccaat	gtggccgaga	ccagcgacgg	1260
gggcgtttac	tgggaaggca	ttgatgagcc	gctagcttca	ggtgtcacca	tcacgtcctg	1320
gaagaataag	gagtggagct	cagaggatgg	ggaaccttgt	gcccaccca	actcgaggtt	1380
ctgcaccct	gccagccagt	gccccatcat	tgatgctgcc	tgggagtctc	cggaagggtg	1440
tcccattgaa	ggcattatct	ttggaggccg	tagacctgct	ggtgtccctc	tagtctatga	1500
agctctcagc	tggcaacatg	gagtctttgt	gggggcggcc	atgagatcag	aggccacagc	1560
ggctgcagaa	cataaaggca	aatcatcat	gcatgacccc	tttgccatgc	ggcccttctt	1620
tggctacaac	ttcggcaa	acctggccca	ctggcttagc	atggcccagc	accagcagc	1680
caaactgccc	aagatcttcc	atgtcaactg	gttccggaag	gacaaggaag	gcaaattcct	1740
ctggccaggc	tttgagaga	actccagggt	gctggagtgg	atgttcaacc	ggatcgatgg	1800
aaaagccagc	accaagctca	cgcccatagg	ctacatcccc	aaggaggatg	ccctgaacct	1860
gaaaggcctg	gggcacatca	acatgatgga	gcttttcagc	atctccaagg	aattctggga	1920
gaaggagggtg	gaagacatcg	agaagtatct	ggaggatcaa	gtcaatgccg	acctcccctg	1980
tgaaatcgag	agagagatcc	ttgccttgaa	gcaaagaata	agccagatgt	aatcagggcc	2040
tgagtgcctt	acctttaaaa	tcattccctt	tcccatccat	aagggtgcagt	aggagcaaga	2100
gagggcaagt	gttcccaa	tgacgccacc	ataataatca	tcaccacacc	gtgagcagat	2160
ctgaaaggca	cactttgatt	tttttaagga	taagaaccac	agaacactgg	gtagtagcta	2220
atgaaattga	gaagggaa	cttagcatgc	ctccaaaaat	tcacatccaa	tgcatagttt	2280
gttcaaattt	aaggttactc	aggcattgat	cttttcagt	ttttttcact	ttagctatgt	2340
ggattagcta	gaatgcacac	caaaaaata	cttgagctgt	atatatatat	gtgtgtgtgt	2400
gtgtgtgtgt	gtgtgtgtgt	gtgtgcatgt	atgtgcacat	gtgtctgtgt	ggtatatattg	2460
tgtatgtgta	tttgtatgta	ctgttattga	aatatatatt	aatacctttg	gaaaaatcct	2520

gggcaagatg acctactagt tttccttgaa aaaaagttgc tttgttatta atattgtgct 2580  
 taaattatTT ttatacacca ttgttcctta cctttacata attgcaatat ttccccctta 2640  
 ctacttcttg gaaaaaaatt acaaaatgaa gttttataga aaagatgg 2688

<210> 18

<211> 3095

<212> DNA

<213> Homo sapiens

<400> 18

tagcagagca atcaccacca agcctggaat aactgcaagg gctctgctga catcttcctg 60  
 aggtgccaaag gaaatgagga tggaggaagg aatgaatgtt ctccatgact ttgggatcca 120  
 gtcaacacat tacctccagg tgaattacca agactcccag gactgggtca tcttggtgtc 180  
 cgtgatcgca gacctcagga atgccttcta cgtcctcttc cccatctggt tccatcttca 240  
 ggaagctgtg ggcattaaac tcctttgggt agctgtgatt ggagactggc tcaacctcgt 300  
 ctttaagtgg atttctctttg gacagcgtcc atactggtgg gttttggata ctgactacta 360  
 cagcaacact tccgtgcccc tgataaagca gttccctgta acctgtgaga ctggaccagg 420  
 gagccccctt ggccatgcca tgggcacagc aggtgtatac tacgtgatgg tcacatctac 480  
 tctttccatc tttcagggaa agataaagcc gacctacaga tttcgggtgct tgaatgtcat 540  
 tttgtggttg ggattctggg ctgtgcagct gaatgtctgt ctgtcacgaa tctaccttgc 600  
 tgctcatTTT cctcatcaag ttgttgctgg agtcctgtca ggcattgctg ttacagaaac 660  
 tttcagccac atccacagca tctataatgc cagcctcaag aaatatTTTc tcattacctt 720  
 cttcctgttc agcttcgcca tcggattTTTa tctgctgctc aagggactgg gtgtagacct 780  
 cctgtggact ctggagaaaag cccagagggtg gtgcgagcag ccagaatggg tccacattga 840  
 caccacaccc tttgccagcc tcctcaagaa cctgggcacg ctctttggcc tggggctggc 900  
 tctcaactcc agcatgtaca gggagagctg caaggggaaa ctcaagcaagt ggctcccatt 960  
 ccgcctcagc tctattgtag cctccctcgt cctcctgcac gtctttgact ccttgaaacc 1020  
 cccatcccaa gtcgagctgg tcttctacgt cttgtccttc tgcaagagtg cggtagtgcc 1080  
 cctggcatcc gtcagtgtca tcccctactg cctcgcccag gtcttgggcc agccgcacaa 1140  
 gaagtcgTTg taagagatgt ggagtcttcg gtgttttaaag tcaacaacca tgccagggat 1200  
 tgaggaggac tactatttga agcaatgggc actgggtattt ggagcaagtg acatgccatc 1260  
 cattctgccg tcgtggaatt aaatcacgga tggcagattg gagggctgcc tggcttattc 1320  
 ccatgtgtga ctccagcctg ccctcagcac agactctttc agatggagggt gccatatcac 1380  
 gtacaccata tgcaagTTTc ccgccaggag gtcctcctct ctctacttga atactctcac 1440  
 aagtagggag ctcaactcca ctggaacagc ccattttatc tttgaatggg cttctgccag 1500

```

cccatTTTga ggccagaggt gctgtcagct caggTggTcc tctTTTaaa tcctaTcat 1560
attgggTaat gTTTTTgaaa agctaTgaa gctattgaga aagacctgtt gctagaagtt 1620
gggtTgttct ggattTTTccc ctgaagactt acttattctt ccgtcacata taaaaagca 1680
agacttccag gtagggccag ctcaaaagcc caggctggag atcctaactg agaattttct 1740
acctgtgttc attcttaccg agaaaaggag aaaggagctc tgaatctgat aggaaaagaa 1800
ggctgcctaa ggaggagttt ttagtatgtg gcgtatcatg caagtgtat gccaagccat 1860
gtctaaatgg cTTTaattat atagtaatgc actctcagta atgggggacc agcttaagta 1920
taattaatag atggTtagtg gggtaatctt gcttctagta TTTTTTTac tTgtcataca 1980
tgttcatcgt atttcctTgg atttctgaat ggctgcagtg acccagatat tgcactaggt 2040
caaaacattc aggtatagct gacatctcct ctatcacatt acatcatcct ccttataagc 2100
ccagctctgc TTTTTccaga ttcttccact ggctccacat ccacccact ggatcttcag 2160
aaggctagag ggcgactctg gtggTgcttt tgtatgtttc aattaggctc tgaaatcttg 2220
ggcaaaatga caaggggagg gccaggattc ctctctcagg tcaactccagt gttactTTta 2280
attcctagag ggtaaatatg actcctttct ctatcccaag ccaaccaaga gcacattctt 2340
aaaggaaaag tcaacatctt ctctctTTTT TTTTTTTTT gagacagggt ctactatgt 2400
tgcccaggct gctcttgaat tcctgggctc aagcagtcct cccaccctac cacagcgtcc 2460
cgcgtagctg gcatacaggt gcaagccact atgtccagct agccaactcc tcctTgcctg 2520
cttttctttt TTTTTctttt ttTgagacgg cgcacctatc acccaggctg gTgtggagtg 2580
gcacgatctt ggctcactgc aacctcttcc tcctggTtca agcgattctc atgtctcagc 2640
ctctcagta gctaggacta ccggcgTgca ccacatgcc aggctaattt ttatattttt 2700
agaatttttag aagagatggg atttcatcat gttggccagg ctggTctcga actcctgacc 2760
tcaagtgatc cacctgcctt ggcctcccaa ggtgctagga ttacaggcat gagccaccgc 2820
accgggccct cctTgcctgt ttttcaatct catctgatat gcagagtatt tctgccccac 2880
ccacctacc cccaaaaaaa gctgaagcct atttatttga aagtcctTgt ttttgctact 2940
aattatatag tataccatac attatcattc aaaacaacca tcctgctcat aacatctttg 3000
aaaagaaaaa tatatatgtg cagtatttta ttaaagcaac attttattta agaataaagt 3060
cttgTtaatt actataTTTT agatgcaatg tgatc 3095

```

&lt;210&gt; 19

&lt;211&gt; 2128

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 19

```

gggggtccca tcgggcccgc cctcgcacgt cactccggga ccccgcggc ctccgcaggt 60

```



tctgcgctcc	aggccggagt	cagagactcc	aggatcggtt	ctttcatctt	cgccgcccct	120
gcgcgtccag	ctcttctaag	acgagatgcc	gtcgggcttc	caacagatag	gctccgaaga	180
tggggaaccc	cctcagcagc	gagtgactgg	gaccctggtc	cttgctgtgt	tctctgcggt	240
gcttggtccc	ctgcagtttg	ggtacaacat	tggggctcatc	aatgcccctc	agaaggatgat	300
tgaacagagc	tacaatgaga	cgtggctggg	gaggcagggg	cctgagggac	ccagctccat	360
ccctccaggc	accctcacca	ccctctgggc	cctctccgtg	gccatctttt	ccgtgggchg	420
catgatttcc	tccttcctca	ttggtatcat	ctctcagtgg	cttggaaagga	aaagggccat	480
gctggtcaac	aatgtcctgg	cggtgctggg	gggcagcctc	atgggctggg	ccaacgctgc	540
tgcttcctat	gaaatgctca	tccttgagcg	attcctcatt	ggcgctact	cagggtgac	600
atcagggtcg	gtgcccattg	acgtggggga	gattgctccc	actcacctgc	ggggcgccct	660
ggggacgctc	aaccaactgg	ccattgttat	cggcattctg	atcgcccagg	tgctgggctt	720
ggagtccctc	ctgggcactg	ccagcctgtg	gccactgctc	ctgggcctca	cagtgtctacc	780
tgccctcctg	cagctggtcc	tgctgccctt	ctgtcccgag	agcccccgct	acctctacat	840
catccagaat	ctcgaggggc	ctgccagaaa	gagtctgaag	cgcttgacag	gctgggcca	900
tgcttctgga	gtgctggctg	agctgaagga	tgagaagcgg	aagctggagc	gtgagcgcc	960
actgtccctg	ctccagctcc	tgggcagccg	taccacccgg	cagcccctga	tcattgcggt	1020
cgtgctgcag	ctgagccagc	agctctctgg	catcaatgct	gttttctatt	attcgaccag	1080
catcttcgag	acagcagggg	taggccagcc	tgccatagcc	accataggag	ctggtgtggt	1140
caacacagtc	ttcaccttgg	tctcgggtgt	gttggtggag	cgggcggggc	gccggacgct	1200
ccatctcctg	ggcctggcgg	gcatgtgtgg	ctgtgccatc	ctgatgactg	tggctctgct	1260
cctgctggag	cgagttccag	ccatgagcta	cgtctccatt	gtggccatct	ttggcttcgt	1320
ggcatttttt	gagattggcc	ctggcccat	tccttggttc	atcgaggccg	agctcttcag	1380
ccagggaccc	cgcccgagc	ccatggctgt	ggctggtttc	tccaactgga	cgagcaactt	1440
catcattggc	atgggtttcc	agtatgttgc	ggaggctatg	gggccctacg	tcttccttct	1500
atttgcggtc	ctcctgctgg	gcttcttcat	cttcaccttc	ttaagagtac	ctgaaactcg	1560
aggccggacg	tttgaccaga	tctcagctgc	cttcacccgg	acaccctctc	ttttagagca	1620
ggagggtgaaa	cccagcacag	aacttgagta	tttagggcca	gatgagaacg	actgaggggc	1680
caggcagggg	tgggagagcc	agctctctct	acccggccca	gagaccctct	cctttcctct	1740
gcagcacttt	aaccctctct	tccctattat	ttccgggtgg	aaaagaatcc	ctgcagcctg	1800
gtagaattgg	gaagctgggg	gaagggtggt	ctgagcacc	cctcattccc	ctcgtgtgac	1860
tctcttggtg	tatttatgtg	ttgtgggttg	gccgtggcca	tcagggtggg	ccactctccc	1920
ctccctcttc	cttcccccat	ccccttccct	ccccaccttc	cccagactca	gctccagaat	1980
accttcttcg	ctgctagaga	agggggattg	gaggggaagac	aggtctagac	tttctcagtg	2040
ggacaaacca	gagcagagag	caggacagga	gacaagaaat	ccagtttccc	accaccttgg	2100

actcctccca caatctggga ctttcact 2128

<210> 20

<211> 1496

<212> DNA

<213> Homo sapiens

<400> 20

ggcacgagga aatcaagat aaaaatgttc acaattaagc tccttctttt tattgttcct	60
ctagtatttt cctccagaat tgatcaagac aattcatcat ttgattctct atctccagag	120
ccaaaatcaa gatttgctat gttagacgat gtaaaaaatt tagccaatgg cctccttcag	180
ttgggacatg gtcttaaaga ctttgtccat aagacgaagg gccaaattaa tgacatat	240
caaaaactca acatatttga tcagtctttt tatgatctat cgctgcaaac cagtgaatc	300
aaagaagaag aaaaggaact gagaagaact acatataaac tacaagtcaa aatgaagag	360
gtaaagaata tgtcacttga actcaactca aaacttgaaa gcctcctaga agaaaaatt	420
ctacttcaac aaaaagtga atatttagaa gagcaactaa ctaacttaat tcaaatcaa	480
cctgaaactc cagaacacc agaagtaact tcacttaaaa cttttgtaga aaaacaagat	540
aatagcatca aagaccttct ccagaccgtg gaagaccaat ataaacaatt aaaccaacag	600
catagtcaaa taaaagaaat agaaaatcag ctccagaagga ctagtattca agaaccaca	660
gaaatttctc tatcttccaa gccaaagagca ccaagaacta ctcccttctc tcagttgaat	720
gaaataagaa atgtaaaaca tgatggcatt cctgctgaat gtaccaccat ttataacaga	780
gggtgaacata caagtggcat gtatgccatc agaccagca actctcaagt tttcatgtc	840
tactgtgatg ttatatcagg tagtccatgg acattaattc aacatcgaat agatggatca	900
caaaacttca atgaaacgtg ggagaactac aaatatggtt ttgggaggct tgatggagaa	960
ttttggttgg gcctagagaa gatatactcc atagtgaagc aatctaatta tgttttacga	1020
attgagttgg aagactggaa agacaacaaa cattatat	1080
aatcacgaaa ccaactatac gctacatcta gttgcgatta ctggcaatgt cccaatgca	1140
atcccggaaa acaaagattt ggtgttttct acttgggatc acaaagcaaa aggacacttc	1200
aactgtccag agggttattc aggaggctgg tgggtggcatg atgagtgtgg agaaaacaac	1260
ctaaatggta aatataacaa accaagagca aaatctaagc cagagaggag aagaggatta	1320
tcttggaggt ctcaaaatgg aaggttatac tctataaaat caacaaaaat gttgatccat	1380
ccaacagatt cagaaagctt tgaatgaact gaggcaatt taaaaggcaa taatttaa	1440
attaacctca ttccaagtta atgtggtcta ataacttggt attaaatcct taagag	1496

<210> 21

<211> 1415

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

```

ggaggagggg gagagagaga agagaagaaa aagaaaaaag aacatcaata aaaagaagtc      60
agatttggttc gaaatcttga ggagtcttca ggccagctcc ctgtcggatg gcttttatga      120
aaaaatatct cctccccatt ctggggctct tcatggccta ctactactat tctgcaaacg      180
aggaattcag accagagatg ctccaaggaa agaaagtgat tgtcacaggg gccagcaaag      240
ggatcggaag agagatggct tatcatctgg cgaagatggg agcccatgtg gtggtgacag      300
cgaggtcaaa agaaactcta cagaagggtg tatccactg cctggagctt ggagcagcct      360
cagcacacta cattgctggc accatggaag acatgacctt cgcagagcaa tttgttgccc      420
aagcagggaaa gctcatggga ggactagaca tgctcattct caaccacatc accaactctt      480
ctttgaatct ttttcatgat gatattcacc atgtgcgcaa aagcatggaa gtcaacttcc      540
tcagttacgt ggtcctgact gtagctgcct tgcccatgct gaagcagagc aatggaagca      600
ttgttgctgt ctcctctctg gctgggaaag tggcttatcc aatggttgct gcctattctg      660
caagcaagtt tgctttggat gggttcttct cctccatcag aaaggaatat tcagtgtcca      720
gggtcaatgt atcaatcact ctctgtgttc ttggcctcat agacacagaa acagccatga      780
aggcagtttc tgggatagtc catatgcaag cagctccaaa ggaggaatgt gccctggaga      840
tcatcaaagg gggagctctg cgccaagaag aagtgtatta tgacagctca ctctggacca      900
ctcttctgat cagaaatcca tgcaggaaga tcctggaatt tctctactca acgagctata      960
atatggacag attcataaac aagtaggaac tccctgaggg ctgggcatgc tgagggatatt 1020
tgggactggt ctgtctcatg tttatctgag ctcttatcta tgaagacatc ttcccagagt 1080
gtccccagag acatgcaagt catgggtcac acctgacaaa tggaaggagt tcctctaaca 1140
tttgcaaaat ggaaatgtaa taataatgaa tgtcatgcac cgctgcagcc agcagttgta 1200
aaattgtttag taaacatagg tataattacc agatagttat attaaattta tatcttatat 1260
ataataatat gtgatgatta atacaatatt aattataata aaggtcacat aaactttata 1320
aattcataac tggtagctat aacttgagct tattcaggat ggtttcttta aaaccataaa 1380
ctgtacaaat gaaatttttc aatatttggt tctta                                1415

```

&lt;210&gt; 22

&lt;211&gt; 1405

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

acaattcaga ggctgctgcc tgcttaggag gttgtagaaa gctctgtagg ttctctctgt 60  
 gtgtcctaca ggagtcttca ggccagctcc ctgtcggatg gcttttatga aaaaatatct 120  
 cctccccatt ctggggctct tcatggccta ctactactat tctgcaaacg aggaattcag 180  
 accagagatg ctccaaggaa agaaagtgat tgtcacaggg gccagcaaag ggatcggaa 240  
 agagatggct tatcatctgg cgaagatggg agcccatgtg gtggtgacag cgagggtcaaa 300  
 agaaactcta cagaaggtgg tatccactg cctggagctt ggagcagcct cagcacacta 360  
 cattgctggc accatggaag acatgacctt cgcagagcaa tttgttgccc aagcaggaaa 420  
 gctcatggga ggactagaca tgctcattct caaccacatc accaactt ctttgaatct 480  
 ttttcatgat gatattcacc atgtgcgcaa aagcatggaa gtcaacttcc tcagttacgt 540  
 ggtcctgact gtagctgcct tgcccatgct gaagcagagc aatggaagca ttgttgctgt 600  
 ctctctctg gctgggaaag tggcttatcc aatggttgct gcctattctg caagcaagtt 660  
 tgctttggat gggttcttct cctccatcag aaaggaatat tcagtgtcca ggggtcaatgt 720  
 atcaatcact ctctgtgttc ttggcctcat agacacagaa acagccatga aggcagtttc 780  
 tgggatagtc catatgcaag cagctccaaa ggaggaatgt gccctggaga tcatcaaagg 840  
 gggagctctg cgccaagaag aagtgtatta tgacagctca ctctggacca ctcttctgat 900  
 cagaaatcca tgcaggaaga tcctggaatt tctctactca acgagctata atatggacag 960  
 attcataaac aagtaggaac tccctgaggg ctgggcatgc tgagggattt tgggactgtt 1020  
 ctgtctcatg tttatctgag ctcttatcta tgaagacatc ttcccagagt gtccccagag 1080  
 acatgcaagt catgggtcac acctgacaaa tggaaggagt tcctctaaca tttgcaaaat 1140  
 ggaaatgtaa taataatgaa tgtcatgcac cgctgcagcc agcagttgta aaattgtag 1200  
 taaacatagg tataattacc agatagttat attaaattta tatcttatat ataataatat 1260  
 gtgatgatta atacaatatt aattataata aaggtcacat aaactttata aattcataac 1320  
 tggtagctat aacttgagct tattcaggat ggtttcttta aaaccataaa ctgtacaaat 1380  
 gaaatttttc aatatttggt tctta 1405

<210> 23

<211> 1944

<212> DNA

<213> Homo sapiens

<400> 23

ccctctcgcg cccagggcgc gtgtaccccc gcactccgcg ccccgcccta gaagctctct 60  
 ctccccgctc cccggcccg ccccgcccc gcccgcccc agcccgctgg gccgcatgg 120  
 agcgctggcc ttggccgctg ggcggcgccct ggctgctcgt ggctgcccgc gcgctgctgc 180  
 agctgctgcg ctgagacctg cgtctgggccc gcccgctgct ggcggcgctg gcgctgctgg 240

ccgcgctcga ctggctgtgc cagcgcctgc tgcccccgcc ggccgcactc gccgtgctgg	300
ccgccgccgg ctggatcgcg ttgtcccgcc tggcgcgccc gcagcgctg ccggtggcca	360
ctcgcgcggt gctcatcacc ggctgtgact ctggtttttg caaggagacg gccaagaaac	420
tggactccat gggcttcacg gtgctggcca ccgtattgga gttgaacagc cccggtgcca	480
tcgagctgcg tacctgctgc tcccctcgcc taaggctgct gcagatggac ctgaccaaac	540
caggagacat tagccgcgtg ctagagttca ccaaggccca caccaccagc accggcctgt	600
ggggcctcgt caacaacgca ggccacaatg aagtagttgc tgatgcggag ctgtctccag	660
tggccacttt ccgtagctgc atggaggtga atttcttttg cgcgctcgag ctgaccaagg	720
gcctcctgcc cctgctgcgc agctcaaggg gccgcacgt gactgtgggg agcccagcgg	780
gggacatgcc atatccgtgc ttgggggcct atggaacctc caaagcggcc gtggcgctac	840
tcatggacac attcagctgt gaactccttc cctggggggt caaggtcagc atcatccagc	900
ctggctgctt caagacagag tcagtgagaa acgtgggtca gtgggaaaag cgcaagcaat	960
tgctgctggc caacctgcct caagagctgc tgcaggccta cggcaaggac tacatcgagc	1020
acttgcatgg gcagttcctg cactcgctac gcctggccat gtccgacctc accccagttg	1080
tagatgccat cacagatgcg ctgctggcag ctcgggcccc cgcgcgctat taccgccggc	1140
agggcctggg gctcatgtac ttcatccact actacctgcc tgaaggcctg cggcgccgct	1200
tcctgcaggc cttcttcac agtcaactgtc tgccctgagc actgcagcct ggccagcctg	1260
gcactacccc accacaggac gcagcccagg gcccaaacct gagccccggc ctttccccag	1320
cagtggctcg gtgagccatg tgcacctatg gccagccac tgcagcacag gaggctccgt	1380
gagcccttgg ttcttccccg aaaacccccca gcattacgat cccccaagtg tcctggaccc	1440
tggcctaaag aatcccaccc ccacttcatg cccactgccg atgcccacatc caggcccgggt	1500
gaggccaagg tttcccagtg agcctctgcg cctctccact gtttcatgag cccaaacacc	1560
ctcctggcac aacgctctac cctgcagctt ggagaactcc gctggatggg gagtctcatg	1620
caagacttca ctgcagcctt tcacaggact ctgcagatag tgccctctgca aactaaggag	1680
tgactaggtg ggttggggac cccctcagga ttgtttctcg gcaccagtgc ctgagtgtg	1740
caattgaggg ctaaattcca agtgtctctt gactggctca agaattaggg ccccaactac	1800
acacccccaa gccacagggg agcatgtact gtacttccca attgccacat tttaaataaa	1860
gacaaatttt tattttcttct aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1920
aaaaaaaaaa aaaaaaaaaa aaaa	1944

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/10036

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 A61K31/517 C07D403/04 A61P3/06

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07D A61K A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, CHEM ABS Data, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE CHEMCATS            CHEMICAL ABSTRACTS SERVICES, COLUMBUS,            OHIO, US;            XP0002225148            ORDER NUMBER: TRG10400#07364-D;            TRG10400#01815-D; TRG10400#01814-D;            TRG10400#01812-D; TRG10400#01811-D;            TRG10400#01809-D; TRG10400#01736-D;            TRG10400#01735-D;            TRG10400#01732-D; TRG10400#01729-D            &amp; "Chem.Folio"            15 May 2001 (2001-05-15), LION BIOSCIENCE            , HEIDELBERG, GERMANY</p> <p style="text-align: center;">-----            -/--</p>	1-3

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \* & \* document member of the same patent family

Date of the actual completion of the international search

12 January 2004

Date of mailing of the international search report

30/01/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Kollmannsberger, M

## INTERNATIONAL SEARCH REPORT

International Publication No

PCT/EP 03/10036

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>COLLINS J L ET AL: "Identification of a Nonsteroidal Liver X Receptor Agonist through Parallel Array Synthesis of Tertiary Amines" JOURNAL OF MEDICINAL CHEMISTRY, AMERICAN CHEMICAL SOCIETY. WASHINGTON, US, vol. 45, 2002, pages 1963-1966, XP002225147 ISSN: 0022-2623 the whole document</p>	1-34
A	<p>GUPTA C M ET AL: "Drugs acting on the central nervous system. Syntheses of substituted quinazolones and quinazolines and triazepino- and triazocinoquinazolones" JOURNAL OF MEDICINAL CHEMISTRY, AMERICAN CHEMICAL SOCIETY. WASHINGTON, US, vol. 11, no. 2, 26 February 1968 (1968-02-26), pages 392-395, XP002156695 ISSN: 0022-2623 example 15; table 1</p>	1-34
A	<p>WO 97/20823 A (CRISCIONE LEOLUCA ; YAMAGUCHI YASUCHIKA (CH); CIBA GEIGY AG (CH); MAH) 12 June 1997 (1997-06-12) page 74; example 28 claims 1,12</p>	1-34

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 03/10036

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
Although claims 13-24, 33,34 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compounds/compositions.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.



### Information on patent family members

PCT/EP 03/10036

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9720823 A	12-06-1997	AU 7692996 A	27-06-1997
		WO 9720823 A2	12-06-1997
		ZA 9610020 A	01-06-1997